UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/728,399	12/05/2003	Jerry R. Colca	01455/1	8615

7590 04/28/2006

Pharmacia Corporation
Global Patent Department
P.O. Box 1027
Mail Zone MC5
St. Louis, MO 63141

EXAMINER	
BOWMAN, AMY HUDSON	

ART UNIT	PAPER NUMBER
1635	

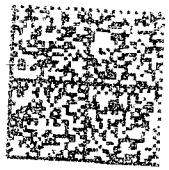
DATE MAILED: 04/28/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

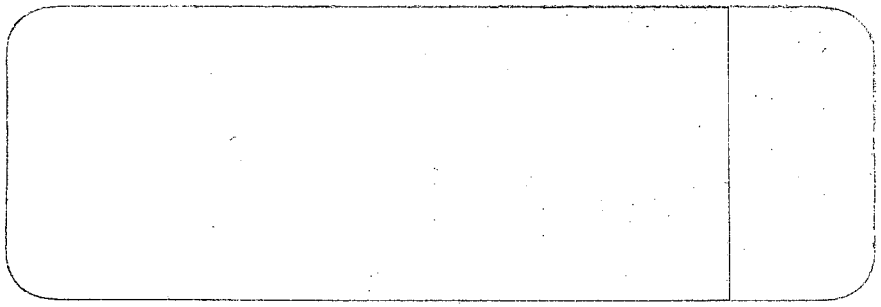
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450
IF UNDELIVERABLE RETURN IN TEN DAYS

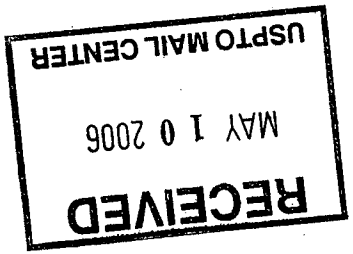
OFFICIAL BUSINESS

AN EQUAL OPPORTUNIT



U.S. OFFICIAL MAIL
PENALTY FOR
PRIVATE USE $300
PITNEY BOWES
02 1A
0004205065
$ 04.05⁰
APR 28 2006
MAILED FROM ZIP CODE 22314





Office Action Summary	Application No.	Applicant(s)	
	10/728,399	COLCA, JERRY R.	
	Examiner	Art Unit	
	Amy H. Bowman	1635	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *09 February 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-25* is/are pending in the application.

4a) Of the above claim(s) *16-25* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *3/3/05, 2/9/06*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Applicant's election with traverse of group I, claims 1-15, and SEQ ID NO: 1, in the reply filed on 2/9/2006 is acknowledged.

Applicant asserts that the examiner has failed to show there is an undue burden. Applicant seems to rely on the fact that the withdrawn process claims that depend from or otherwise include all the limitations of the allowable product claim will be rejoined in accordance with the provisions of MPEP § 821.04 to establish that there is not an undue search burden.

Contrary to applicant's assertion, until an elected product claim is found allowable, an otherwise proper restriction requirement between product claims and process claims may be maintained, as explained in the office action mailed 5/12/2005. See "Guidance on Treatment of Product and Process Claims in light of *In re Ochiai, In re Brouwer* and 35 U.S.C. § 103(b)," 1184 O.G. 86 (March 26, 1996) for further explanation of product and process claims. Simply because the process claims may be rejoined upon determination that the product claims are allowable does not mean that there is not a search burden to examine the two groups together before the product claims are found allowable.

The requirement for restriction is still deemed proper and is therefore made FINAL.

Claims 16-25 are withdrawn from further consideration pursuant to 37 CFR 1.142(b), as being drawn to nonelected inventions, there being no allowable generic or linking claim. Additionally, the subject matter of claims 3-6 that is not drawn to the

elected invention, SEQ ID NO: 1, has been withdrawn as being drawn to nonelected inventions. Applicant timely traversed the restriction (election) requirement in the reply filed on 2/9/2006.

Claim Rejections - 35 USC § 112

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claim 5 is rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 5 recites, "consists of at least 8 contiguous..." The terminology "consists of" constitutes closed language, so the language "consists of" coupled with "at least" is contradictory language within the claim . Appropriate correction is required.

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claims 1-15 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

The invention of the above claims is drawn to an antisense compound targeted to a nucleic acid molecule encoding mitoNEET, wherein said antisense compound specifically hybridizes with and inhibits the expression of mitoNEET.

At the outset, it is noted that the claims do not recite a specific target nucleotide sequence by SEQ ID NO, but rather refers to the broad genus of mitoNEET sequences.

The claims encompass any antisense oligonucleotide 8 to 30 nucleobases in length targeted to any nucleic acid molecule encoding any mitoNEET. The claims encompass antisense oligonucleotides targeting any mitoNEET RNA, as well as any mitoNEET homolog or allele known or yet to be discovered from any species of mitoNEET, as well as DNA genomic fragments, splice variants or polynucleotide fragments that express proteins that retain mitoNEET -like activity.

The specification refers to "mitoNEET" as a family of polypeptides (see page 1 of the instant specification), rather than a specific target sequence. The instant specification does not teach antisense oligonucleotides targeted to a family of polypeptides, but rather to the human mitoNEET sequence disclosed as SEQ ID NO: 624. Although the specification discloses specific oligonucleotide sequences targeted to instant SEQ ID NO: 624, the specification does not describe oligonucleotides directed to any other species or variant of mitoNEET to describe the instantly claimed genus of the broad claims which encompass any mitoNEET target sequence. It is the structure of each specific oligonucleotide that leads to its function with regards to a specific target sequence. One of ordinary skill in the art could not make such oligos targeted to any mitoNEET without knowledge of the sequence. Given the breadth of sequences

embraced in the instantly claimed genus, one could not envision the member oligonucleotides that target such a broad genus.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-15 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a specific and substantial utility or a well established utility.

The invention of the above claims is drawn to an antisense compound targeted to a nucleic acid molecule encoding mitoNEET, wherein said antisense compound specifically hybridizes with and inhibits the expression of mitoNEET. The compound is further specified to be an antisense oligonucleotide, more specifically SEQ ID NO: 1. The claims are further drawn to a composition comprising the antisense oligonucleotide and a pharmaceutically acceptable carrier or diluent.

The specification discloses asserted utilities for mitoNEET and antisense oligonucleotides targeted to mitoNEET. The specification discloses that mitoNEET is likely to have therapeutic uses in a variety of cardiovascular, endothelial, and angiogenic disorders, including systemic disorders that affect vessels, such as diabetes mellitus (see paragraph 30 of the instant specification). The instant specification discloses a multitude of diseases or disorders that mitoNEET could have a therapeutic utility with. The specification further discloses that mitoNEET or modulators thereof are

likely to play an important role in the etiology and pathogenesis of many or all of the disorders noted in the specification (see paragraph 63 of the instant specification).

However, hypothesizing various potential utilities for inhibiting mitoNEET with an antisense oligonucleotide is not sufficient for establishing an actual utility specific for mitoNEET inhibition.

Although the instant specification teaches antisense inhibition of mitoNEET expression, the instant specification only offers support for prophetic examples of what utility there is in inhibiting mitoNEET with an antisense oligonucleotide of the invention.

The specification does not disclose a nexus between any specific disease states and a decrease in the expression of mitoNEET. This deficiency is not compensated for in the art. As evidenced by applicant's own post-filing publication, Colca et al. (Am J Physiol Endocrinol Metab, 2004, 286, pages E252-E260), significant further research would have to be conducted to identify such a nexus. Colca et al. teach that mitoNEET is a novel mitochondrial protein that binds to insulin sensitizing, antidiabetic thiazolodinediones and that mitoNEET is a target for thiazolodinediones. Colca et al. teach that the mechanism of thiazolodinediones is not clearly understood. Colca et al. teach that thiazolodinediones may produce their effect by direct interaction with PPARγ, but that not all PPARγ activators have antidiabetic actions. Colca et al. teach that thiazolodinediones may have different effects in different tissues and that resolution of the key site of thiazolodinedione interaction would undoubtedly help in the production of improved therapeutic agents. Colca et al. also teach that mitoNEET may play a role in regulating mitochondrial oxidation of fatty acids, and modulation of this target may be

involved in the mechanism of action of these drugs. Colca et al. conclude that this hypothesis needs to be tested. The teachings of Colca et al. support that there is not an established utility for mitoNEET, but rather the utility of inhibiting mitoNEET is still being explored. Applicant has offered a multitude of asserted utilities for mitoNEET, but has not taught any specific utility that is not simply prophetic.

Therefore, it is unclear what specific and substantial utility there is in inhibiting mitoNEET with an antisense oligonucleotide of the instant invention.

Claims 1-15 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Claim Rejections - 35 USC § 102 or 103

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1, 2, 3, and 5 are rejected under 35 U.S.C. 102(b) as anticipated by or, in the alternative, under 35 U.S.C. 103(a) as obvious over Chew et al. (WO 01/79220 A2).

The invention of the above claims is drawn to an antisense compound targeted to a nucleic acid molecule encoding mitoNEET, wherein said antisense compound specifically hybridizes with and inhibits the expression of mitoNEET. The compound is further specified to be an antisense oligonucleotide comprising at least 8 contiguous nucleic acids of a nucleic acid sequence of SEQ ID NO: 1.

Chew et al. disclose a 10 nucleobase oligonucleotide that comprises 8 contiguous nucleic acids of a nucleic acid sequence of instant SEQ ID NO: 1 (see nucleobases 1-8 of SEQ ID NO: 39 of Chew et al.).

The oligonucleotide taught by Chew et al. is an oligonucleotide primer. However, the instant specification does not define an "antisense oligonucleotide" to exclude primers. The instant specification, page 10, defines the term "oligonucleotide" to refer to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof. Furthermore, the oligonucleotide primer taught by Chew et al. is complementary, or antisense, to the target nucleic acid. Therefore, the oligonucleotide primer taught by Chew et al. meets the instant limitation of an antisense oligonucleotide.

The sequence disclosed by Chew et al. meets all of the structural requirements of the instant claims, so the oligonucleotide would also be expected to target a nucleic acid molecule encoding mitoNEET, as well as specifically hybridize and inhibit the expression of mitoNEET, as instantly claimed, absent evidence to the contrary.

See, for example, MPEP § 2112, which states "[w]here applicant claims a composition in terms of a function, property or characteristic and the composition of the prior art is the same as that of the claim but the function is not explicitly disclosed by the reference, the examiner may make a rejection under both 35 U.S.C. 102 and 103, expressed as a 102/103 rejection. 'There is nothing inconsistent in concurrent rejections for obviousness under 35 U.S.C. 103 and for anticipation under 35 U.S.C. 102.' In re Best, 562 F.2d 1252, 1255 n.4, 195 USPQ 430, 433 n.4 (CCPA 1977). This same rationale should also apply to product, apparatus, and process claims claimed in terms of function, property or characteristic. Therefore, a 35 U.S.C. 102/103 rejection is appropriate for these types of claims as well as for composition claims."

Therefore, the instant invention is anticipated or obvious over Chew et al.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Amy H. Bowman whose telephone number is 571-272-0755.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Andrew Wang can be reached on 571-272-0811. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Patent applicants with problems or questions regarding electronic images that can be viewed in the Patent Application Information Retrieval system (PAIR) can now contact the USPTO's Patent Electronic Business Center (Patent EBC) for assistance.

Representatives are available to answer your questions daily from 6 am to midnight (EST). The toll free number is (866) 217-9197. When calling please have your application serial or patent number, the type of document you are having an image problem with, the number of pages and the specific nature of the problem. The Patent Electronic Business Center will notify applicants of the resolution of the problem within 5-7 business days. Applicants can also check PAIR to confirm that the problem has been corrected. The USPTO's Patent Electronic Business Center is a complete service center supporting all patent business on the Internet. The USPTO's PAIR system provides Internet-based access to patent application status and history information. It also enables applicants to view the scanned images of their own application file folder(s) as well as general patent information available to the public. For more information about the PAIR system, see http://pair-direct.uspto.gov.

For all other customer support, please call the USPTO Call Center (UCC) at 800-786-9199.

Amy H. Bowman
Examiner
Art Unit 1635

JAMES SCHULTZ, PH.D.
PRIMARY EXAMINER

FORM PTO-1449 U.S. DEPARTMENT OF COMMERCE, PATENT AND TRADEMARK OFFICE	ATTY. DOCKET NO: PC027829A	APPLICATION NO.: 10/728,399
SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT	INVENTOR: Jerry R. Colca	
(Use several sheets if necessary)	Filed: 12/5/03	Group: 1635

U.S. PATENT DOCUMENTS

Examiner Initial		Document Number	Date	Name	Class	Subclass	Filing Date If Appropriate
CMB		US 5,801,154A	9/1/98	Baracchini, et al.			
CMB		US5,998,148A	12/7/99	Bennett, et al.			

FOREIGN PATENT DOCUMENTS

Examiner Initial		Document Number	Date	Country	Class	Subclass	Translation Yes	Translation No
CMB	B1	WO 03/031571 A2	4/17/03	WO	C12N		✓	

EXAMINER CMB	DATE CONSIDERED 4/26/06

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

FORM PTO-1449 U.S. DEPARTMENT OF COMMERCE, PATENT AND TRADEMARK OFFICE	ATTY. DOCKET NO: PC027829A	APPLICATION NO.: 10/728,399
SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT	INVENTOR: Jerry R. Colca	
(Use several sheets if necessary)	Filed: 12/5/03	Group: 1635

OTHER DOCUMENTS (Including Author, Title, Date, Pertinent Pages, Etc.)

Examiner Initial		
ATH	C1	Collins, F. S., *Generation and Initial Analysis of More than 15,000 Full-Length Human and Mouse CDNA Sequences*, December 24, 2002, Vol. 99, No. 26, pgs. 16899-16903
	C2	Taylor, M. F., et al., *Antisense Oligonucleotides: A Systematic High-Throughput Approach to Target Validation and Gene Function Determination*, DDT. , December 1999, Vol. 4, No. 12, pgs. 562-567
	C3	Jen, K, et al., *Suppression of Gene Expression by Targeted Disruption of Messenger RNA: Available Options and Current Strategies*. Stem Cells, 2000, Vol. 18, pgs. 307-319
	C4	Branch, A.D., *A Good Antisense Molecule is Hard to Find*, TIBS, February 1998, pgs. 45-50
	C5	Green, D. W., et al., *Antisense Oligonucleotides: An Evolving Technology for the Modulation of Gene Expression in Human Disease*, J. Am. Coll. Surg., Vol. 191, No. 1, July 2000, pgs. 93-105

EXAMINER ATH	DATE CONSIDERED 4/26/06

EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

INFORMATION DISCLOSURE CITATION *(Use several sheets if necessary)*	ATTY. DOCKET NO. PC27829A	SERIAL NO. 10/728,399
	APPLICANT Jerry R. Colca	
	FILING DATE December 5, 2003	GROUP ~~1636~~ 1625

U.S. PATENT DOCUMENTS

EXAMINER INITIAL	DOCUMENT NUMBER	DATE	NAME	CLASS	SUBCLASS	FILING DATE IF APPROPRIATE

FOREIGN PATENT DOCUMENTS

	DOCUMENT NUMBER	DATE	COUNTRY	CLASS	SUBCLASS	TRANSLATION YES	TRANSLATION NO
AHM	WO 02436716	11/29/01	International	A61K	31/00		
AHM	WO 94250269	4/25/94	International	A61K	31/42		

OTHER DOCUMENTS *(Including Author, Title, Date, Pertinent Pages, Etc.)*

AHM			Chang, et al., Diabetes, "Ciglitazone, a new hypoglycemic agent. II. Effect on glucose and lipid metabolisms and insulin binding in the adipose tissue of C57BL/6J-ob/ob and -+/? Mice" , Vol. 32, pages 839-845 (1983)
			Colca and Morton New Antidiabetic Drugs, "Antihyperglycaemic Thiazolidinediones: Ciglitazone and Its Analogues", C.J. Bailey and P.R. Flatt, eds., Smith-Gordon, New York, pages 255-261 (1990)
			Dreyer, C., et al., Cell, "Control of the peroxisomal β-oxidation pathway by a novel family of nuclear hormone receptors", Vol. 68, pages 879-887 (1992)
			Goke, R., et al., Digestion, "Pioglitazone Inhibits Growth of Carcinoid Cells and Promotes TRAIL-Induced Apoptosis by Induction of p21", Vol. 64(2), pages 75-80 (2001)
			Gottlicher, Proc. Natl. Acad., "Fatty acids activate a chimera of the clofibric acid-activated receptor and the glucocorticoid receptor", Vol. 89, pages 4653-4657 (1992)
			Harris and Kletzien, Molecular Pharmacology, "Localization of a pioglitazone response element in the adipocyte fatty

AHM 4/26/06

00INFORMATION DISCLOSURE CITATION *(Use several sheets if necessary)*	ATTY. DOCKET NO. PC27829A	SERIAL NO. 10/728,399
	APPLICANT Jerry R. Colca	
	FILING DATE December 5, 2003	GROUP ~~1636~~ 1615

			acid-binding protein gene", Vol. 45, pages 439-445 (1994)
			Hiragum, et al., Journal Cell Physiology, "Preadipocyte Differentiation in Vitro: Identification of a Highly Active Adipogenic Agent:", Vol. 134, pages 124-130 (1988)
			Isseman and Green, Nature, "Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators" Vol. 347, pages 645-650 (1990
			Itami, A., et al., International Journal of Cancer, "Ligands for peroxisome proliferator-activated receptor γ inhibit growth of pancreatic cancers both *in vitro* and *in vivo*", Vol. 94(3), pages 370-376 (2001)
			Keller and Whali, Trends Endocrinolgy Met., Vol. 4, pages 291-296 (1993)
			Kletzien, et al., Mol. Pharmacology, "Enhancement of adipocyte differentiation by an insulin-sensitizing agent", Vol. 41, pages 393-398 (1992)
			Kliewer, et al., Proc. Natl. Acad. Sci., "Differential Expression and Activation of a Family of Murine Peroxisome Proliferator-Activated Receptors", Vol. 91, pages 7355-7359 (1994)
			Mueller, E., et al., Molecular Cell, "Terminal Differentiation of Human Breast Cancer through PPARγ", Vol. 1(3), pages 465-470 (1998)
			Nolan, et al., New England Journal of Medicine, "Improvement in Glucose Tolerance and Insulin Resistance in Obese Subjects Treated with Troglitazone", Vol. 331, pages 1188-1193 (1994)
			Okano, H., et al., Anti-Cancer Drugs, "Peroxisome proliferator-activated receptor *[gamma]* augments tumor necrosis factor family-induced apoptosis in hepatocellular carcinoma", Vol. 13(1), pages 59-65 (2002)
			Sparks, et al., Journal Cell Physiology, "Antidiabetic AD4743 Enhances Adipocyte Differentiation of 3T3 Mesenchymal Stem Cells", Vol. 146, pages 101-109 (1991)
			Tankaka, T., et al., Cancer Research, "Ligands for Peroxisome Proliferator-activated Receptors α and γ Inhibit Chemically Induced Colitis and Formation of Aberrant Crypt Foci in Rats" Vol. 61(6), pages 2424-2428 (2001)
			Tontonoz, et al., Genes and Development, "mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer" Vol. 8, pages 1224-1234 (1994)
			Williams, et al., Diabetes, Vol. 42, page 59A (1993)
			Zhu, et al., "Cloning of a new member of the peroxisome proliferator-activated receptor gene family from mouse liver" J. Biol. Chem., Vol. 268, pages 26817-26820 (1993)
			Genbank: NM 018464
			Genbank: NM 134007

EXAMINER [signature] DATE CONSIDERED 4/26/06

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

Notice of References Cited	Application/Control No. 10/728,399	Applicant(s)/Patent Under Reexamination COLCA, JERRY R.	
	Examiner Amy H. Bowman	Art Unit 1635	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N	WO 01/79220 A2	10-2001	WO	CHEW et al.	514/44
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Colca et al., Identification of a novel mitochondrial protein ("mitoNEET") cross-linked specifically by a thiazolidinedione photoprobe, 2004, Am J Physiol Endocrinol Metab, 286, pages E252-E260.
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79220 A2

(51) International Patent Classification[7]: C07H

(21) International Application Number: PCT/US01/11941

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/196,406 12 April 2000 (12.04.2000) US

(71) Applicant *(for all designated States except US)*: **GENAISSANCE PHARMACEUTICALS, INC.** [US/US]; Five Science Park, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CHEW, Anne** [US/US]; 1477 Beacon Street #64, Brookline, MA 02446 (US). **CHOI, Julie, Y.** [US/US]; 38 Elizabeth Street, West Haven, CT 06516 (US). **KOSHY, Beena** [IN/US]; Apt. 11B, 1298 Hartford Turnpike, North Haven, CT 06473 (US).

(74) Agent: **SHANER, Sandra, L.**; Genaissance Pharmaceuticals, Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

1-88

WO 01/79220 A2

(54) Title: HAPLOTYPES OF THE HRH2 GENE

(57) Abstract: Novel single nucleotide polymorphisms in the human Histamine H2 receptor (HRH2) gene are described. In addition, various genotypes, haplotypes and haplotype pairs for the HRH2 gene that exist in the population are described. Compositions and methods for haplotyping and/or genotyping the HRH2 gene in an individual are also disclosed. Polynucleotides containing one or more of the HRH2 polymorphisms disclosed herein are also described.

HAPLOTYPES OF THE HRH2 GENE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/196,406 filed April 12, 2000.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically-important proteins. In particular, this invention provides genetic variants of the human Histamine H2 receptor (HRH2) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a lead compound that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by activity at non-intended targets. The lead compound identified in this screening process then undergoes further *in vitro* and *in vivo* testing to determine its absorption, disposition, metabolism and toxicological profiles. Typically, this testing involves use of cell lines and animal models with limited, if any, genetic diversity.

What this approach fails to consider, however, is that natural genetic variability exists between individuals in any and every population with respect to pharmaceutically-important proteins, including the protein targets of candidate drugs, the enzymes that metabolize these drugs and the proteins whose activity is modulated by such drug targets. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a pharmaceutically-important protein may be manifested as significant variation in expression, structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in the treatment of individuals with a drug whose design is based upon a single representative example of the target or enzyme(s) involved in metabolizing the drug. For example, it is well-established that some drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. Also, there is significant variation in how well people metabolize drugs and other exogenous chemicals, resulting in substantial interindividual variation in the toxicity and/or efficacy of such exogenous substances (Evans et al., 1999, *Science* 286:487-491). This variability in efficacy or toxicity of a drug in genetically-diverse patients makes many drugs ineffective or even dangerous in certain groups of the population, leading to the failure of such drugs in clinical trials or their early withdrawal from the market even though they could be highly beneficial

for other groups in the population. This problem significantly increases the time and cost of drug discovery and development, which is a matter of great public concern.

It is well-recognized by pharmaceutical scientists that considering the impact of the genetic variability of pharmaceutically-important proteins in the early phases of drug discovery and development is likely to reduce the failure rate of candidate and approved drugs (Marshall A 1997 *Nature Biotech* **15**:1249-52; Kleyn PW et al. 1998 *Science* **281**: 1820-21; Kola I 1999 *Curr Opin Biotech* **10**:589-92; Hill AVS et al. 1999 in *Evolution in Health and Disease* Stearns SS (Ed.) Oxford University Press, New York, pp 62-76; Meyer U.A. 1999 in *Evolution in Health and Disease* Stearns SS (Ed.) Oxford University Press, New York, pp 41-49; Kalow W et al. 1999 *Clin. Pharm. Therap.* **66**:445-7; Marshall, E 1999 *Science* **284**:406-7; Judson R et al. 2000 *Pharmacogenomics* **1**:1-12; Roses AD 2000 *Nature* **405**:857-65). However, in practice this has been difficult to do, in large part because of the time and cost required for discovering the amount of genetic variation that exists in the population (Chakravarti A 1998 *Nature Genet* **19**:216-7; Wang DG et al 1998 *Science* **280**:1077-82; Chakravarti A 1999 *Nat Genet* **21**:56-60 (suppl); Stephens JC 1999 *Mol. Diagnosis* **4**:309-317; Kwok PY and Gu S 1999 *Mol. Med. Today* **5**:538-43; Davidson S 2000 *Nature Biotech* **18**:1134-5).

The standard for measuring genetic variation among individuals is the haplotype, which is the ordered combination of polymorphisms in the sequence of each form of a gene that exists in the population. Because haplotypes represent the variation across each form of a gene, they provide a more accurate and reliable measurement of genetic variation than individual polymorphisms. For example, while specific variations in gene sequences have been associated with a particular phenotype such as disease susceptibility (Roses AD *supra*; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74) and drug response (Wolfe CR et al. 2000 *BMJ* **320**:987-90; Dahl BS 1997 *Acta Psychiatr Scand* **96** (Suppl 391): 14-21), in many other cases an individual polymorphism may be found in a variety of genomic backgrounds, i.e., different haplotypes, and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *supra*; Drysdale et al. 2000 *PNAS* **97**:10483-10488). Thus, there is an unmet need in the pharmaceutical industry for information on what haplotypes exist in the population for pharmaceutically-important genes. Such haplotype information would be useful in improving the efficiency and output of several steps in the drug discovery and development process, including target validation, identifying lead compounds, and early phase clinical trials (Marshall et al., *supra*).

One pharmaceutically-important gene for the treatment of acid-peptic disorders of the gastrointestinal tract is the Histamine H2 receptor (HRH2) gene or its encoded product. Histamine is a powerful gastric secretagogue and evokes a copious secretion of acid from parietal cells by acting on HRH2 (Babe, K and Serafin, W *The Pharmacological Basis of Therapeutics In. Histamine, Bradykinin and Their Antagonists* 96, pp. 581- 600). The physiological effects of histamine are mediated by HRH2 linked to adenylate cyclase activation and cellular cAMP generation. HRH2 is one of three

subclasses of histamine receptors and plays an important role mediates gastric acid secretion, and antagonists for this receptor have proven to be effective therapy for peptic acid disorders of the gastrointestinal tract (Gantz et al., *Proc. Natl. Acad. Sci. U. S. A* 1991; 88:429-433). HRH2 antagonists, such as cimetidine, ranitidine, and famotidine, competitively inhibit the interaction of histamine with HRH2. They inhibit gastric acid secretion elicited by histamine, as well as by gastric and some muscarinic agonists (Babe, K and Serafin, W, *supra*).

HRH2 is also involved in regulating intestinal secretion (Del Valle and Gantz, *Am. J. Physiol* 1997; 273:G987-G996). Histamine was observed to inhibit prostaglandin E2-stimulated duodenal epithelial bicarbonate secretion through an HRH2 located on enteric neurons. This observation adds another pathway (in addition to stimulating gastric acid secretion) through which histamine can induce duodenal mucosal damage. Neuronal histamine also appears to be involved in the regulation of colonic secretion. Several studies revealed that histamine-mediated activation of HRH2 led to an increase in short-circuit current and Cl secretion by stimulating cholinergic neurons that utilize muscarinic and nicotinic synapses and by activating vasoactive intestinal polypeptide (VIP)-ergic pathways. Together, these findings may in part explain the motility and secretory abnormalities associated with intestinal inflammation (Del Valle and Gantz, *supra*).

An intriguing and potentially exciting role for HRH2 has been in the regulation of cell proliferation (Del Valle and Gantz, *supra*). Human mammary and gastric carcinoma cells and several human melanoma cell lines express HRH2. More recently, it was demonstrated that histamine could stimulate the proliferation of two gastric carcinoma cell lines through interactions with HRH2. The interesting observation that cimetidine improved survival in patients with gastric cancer adds greater potential significance to the role of histamine in cell growth (Del Valle and Gantz, *supra*).

These results indicate that HRH2 not only regulates a limited number of targets such as contractile activity of the heart and gastric acid secretion, but a broader range of biological actions extending from cell differentiation and proliferation to gastrointestinal motility are also sites of HRH2 activity.

The Histamine H2 receptor gene is located on chromosome 5 and contains 1 exon that encodes a 359 amino acid protein. A reference sequence for the HRH2 gene is shown in Figure 1 (GenBank Accession No. AB023486.1; SEQ ID NO:1). Reference sequences for the coding sequence (GenBankAccession No. D49783.1) and protein are shown in Figures 2 (SEQ ID NO:2) and 3 (SEQ ID NO:3), respectively.

Because of the potential for variation in the HRH2 gene to affect the expression and function of the encoded protein, it would be useful to know whether polymorphisms exist in the HRH2 gene, as well as how such polymorphisms are combined in different copies of the gene. Such information could be applied for studying the biological function of HRH2 as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 10 novel polymorphic sites in the HRH2 gene. These polymorphic sites (PS) correspond to the following nucleotide positions in the indicated GenBank Accession Number: 11881 (PS1), 12001 (PS2), 12010 (PS3), 12352 (PS4), 12502 (PS5), 12649 (PS6), 12850 (PS7), 13104 (PS8), 13416 (PS9) and 13497 (PS10) in AB023486.1. The polymorphisms at these sites are guanine or adenine at PS1, guanine or adenine at PS2, guanine or adenine at PS3, cytosine or thymine at PS4, guanine or adenine at PS5, cytosine or thymine at PS6, guanine or adenine at PS7, adenine or guanine at PS8, guanine or cytosine at PS9 and cytosine or thymine at PS10. In addition, the inventors have determined the identity of the alleles at these sites in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African descent, Asian, Caucasian and Hispanic/Latino. From this information, the inventors deduced a set of haplotypes and haplotype pairs for PS1-PS10 in the HRH2 gene, which are shown below in Tables 4 and 3, respectively. Each of these HRH2 haplotypes defines a naturally-occurring isoform (also referred to herein as an "isogene") of the HRH2 gene that exists in the human population.

Thus, in one embodiment, the invention provides a method, composition and kit for genotyping the HRH2 gene in an individual. The genotyping method comprises identifying the nucleotide pair that is present at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10 in both copies of the HRH2 gene from the individual. A genotyping composition of the invention comprises an oligonucleotide probe or primer which is designed to specifically hybridize to a target region containing, or adjacent to, one of these novel HRH2 polymorphic sites. A genotyping kit of the invention comprises a set of oligonucleotides designed to genotype each of these novel HRH2 polymorphic sites. The genotyping method, composition, and kit are useful in determining whether an individual has one of the haplotypes in Table 4 below or has one of the haplotype pairs in Table 3 below.

The invention also provides a method for haplotyping the HRH2 gene in an individual. In one embodiment, the haplotyping method comprises determining, for one copy of the HRH2 gene, the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10. In another embodiment, the haplotyping method comprises determining whether one copy of the individual's HRH2 gene is defined by one of the HRH2 haplotypes shown in Table 4, below, or a sub-haplotype thereof. In a preferred embodiment, the haplotyping method comprises determining whether both copies of the individual's HRH2 gene are defined by one of the HRH2 haplotype pairs shown in Table 3 below, or a sub-haplotype pair thereof. The method for establishing the HRH2 haplotype or haplotype pair of an individual is useful for improving the efficiency and reliability of several steps in the discovery and development of drugs for treating diseases associated with HRH2 activity, e.g., acid-peptic disorders of the gastrointestinal tract.

For example, the haplotyping method can be used by the pharmaceutical research scientist to

validate HRH2 as a candidate target for treating a specific condition or disease predicted to be associated with HRH2 activity. Determining for a particular population the frequency of one or more of the individual HRH2 haplotypes or haplotype pairs described herein will facilitate a decision on whether to pursue HRH2 as a target for treating the specific disease of interest. In particular, if variable HRH2 activity is associated with the disease, then one or more HRH2 haplotypes or haplotype pairs will be found at a higher frequency in disease cohorts than in appropriately genetically matched controls. Conversely, if each of the observed HRH2 haplotypes are of similar frequencies in the disease and control groups, then it may be inferred that variable HRH2 activity has little, if any, involvement with that disease. In either case, the pharmaceutical research scientist can, without *a priori* knowledge as to the phenotypic effect of any HRH2 haplotype or haplotype pair, apply the information derived from detecting HRH2 haplotypes in an individual to decide whether modulating HRH2 activity would be useful in treating the disease.

The claimed invention is also useful in screening for compounds targeting HRH2 to treat a specific condition or disease predicted to be associated with HRH2 activity. For example, detecting which of the HRH2 haplotypes or haplotype pairs disclosed herein are present in individual members of a population with the specific disease of interest enables the pharmaceutical scientist to screen for a compound(s) that displays the highest desired agonist or antagonist activity for each of the most frequent HRH2 isoforms present in the disease population. Thus, without requiring any *a priori* knowledge of the phenotypic effect of any particular HRH2 haplotype or haplotype pair, the claimed haplotyping method provides the scientist with a tool to identify lead compounds that are more likely to show efficacy in clinical trials.

The method for haplotyping the HRH2 gene in an individual is also useful in the design of clinical trials of candidate drugs for treating a specific condition or disease predicted to be associated with HRH2 activity. For example, instead of randomly assigning patients with the disease of interest to the treatment or control group as is typically done now, determining which of the HRH2 haplotype(s) disclosed herein are present in individual patients enables the pharmaceutical scientist to distribute HRH2 haplotypes and/or haplotype pairs evenly to treatment and control groups, thereby reducing the potential for bias in the results that could be introduced by a larger frequency of a HRH2 haplotype or haplotype pair that had a previously unknown association with response to the drug being studied in the trial. Thus, by practicing the claimed invention, the scientist can more confidently rely on the information learned from the trial, without first determining the phenotypic effect of any HRH2 haplotype or haplotype pair.

In another embodiment, the invention provides a method for identifying an association between a trait and a HRH2 genotype, haplotype, or haplotype pair for one or more of the novel polymorphic sites described herein. The method comprises comparing the frequency of the HRH2 genotype, haplotype, or haplotype pair in a population exhibiting the trait with the frequency of the HRH2 genotype, haplotype, or haplotype pair in a reference population. A higher frequency of the

HRH2 genotype, haplotype, or haplotype pair in the trait population than in the reference population indicates the trait is associated with the HRH2 genotype, haplotype, or haplotype pair. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. In a particularly preferred embodiment, the HRH2 haplotype is selected from the haplotypes shown in Table 4, or a sub-haplotype thereof. Such methods have applicability in developing diagnostic tests and therapeutic treatments for acid-peptic disorders of the gastrointestinal tract.

In yet another embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the HRH2 gene or a fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at PS1, adenine at PS2, adenine at PS3, thymine at PS4, adenine at PS5, thymine at PS6, adenine at PS7, guanine at PS8, cytosine at PS9 and thymine at PS10.

A particularly preferred polymorphic variant is an isogene of the HRH2 gene. A HRH2 isogene of the invention comprises guanine or adenine at PS1, guanine or adenine at PS2, guanine or adenine at PS3, cytosine or thymine at PS4, guanine or adenine at PS5, cytosine or thymine at PS6, guanine or adenine at PS7, adenine or guanine at PS8, guanine or cytosine at PS9 and cytosine or thymine at PS10. The invention also provides a collection of HRH2 isogenes, referred to herein as a HRH2 genome anthology.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for a HRH2 cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig.2) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 45, adenine at a position corresponding to nucleotide 195, thymine at a position corresponding to nucleotide 342, adenine at a position corresponding to nucleotide 543 and guanine at a position corresponding to nucleotide 797. A particularly preferred polymorphic cDNA variant comprises the coding sequence of a HRH2 isogene defined by haplotypes 2-11.

Polynucleotides complementary to these HRH2 genomic and cDNA variants are also provided by the invention. It is believed that polymorphic variants of the HRH2-gene will be useful in studying the expression and function of HRH2, and in expressing HRH2 protein for use in screening for candidate drugs to treat diseases related to HRH2 activity.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express HRH2 for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the HRH2 protein. The reference amino acid sequence

comprises SEQ ID NO:3 (Fig.3) and the polymorphic variant comprises serine at a position corresponding to amino acid position 266. A polymorphic variant of HRH2 is useful in studying the effect of the variation on the biological activity of HRH2 as well as on the binding affinity of candidate drugs targeting HRH2 for the treatment of acid-peptic disorders of the gastrointestinal tract.

The present invention also provides antibodies that recognize and bind to the above polymorphic HRH2 protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

The present invention also provides nonhuman transgenic animals comprising one of the HRH2 polymorphic genomic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the HRH2 isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against HRH2 protein, and for testing the efficacy of therapeutic agents and compounds for acid-peptic disorders of the gastrointestinal tract in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the HRH2 gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the HRH2 gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing HRH2 haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the HRH2 gene (Genbank Accession Number AB023486.1; contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence indicated with a bracket ([or]) and the numerical position below the sequence and the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence. SEQ ID NO:54 is equivalent to Figure 1, with the two alternative allelic variants of each polymorphic site indicated by the appropriate nucleotide symbol (R = G or A, Y = T or C, M = A or C, K = G or T, S = G or C, and W = A or T; WIPO standard ST.25).

Figure 2 illustrates a reference sequence for the HRH2 coding sequence (contiguous lines; SEQ ID NO:2) with the polymorphic site(s) and polymorphism(s) identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3 illustrates a reference sequence for the HRH2 protein (contiguous lines; SEQ ID NO:3), with the variant amino acid(s) caused by the polymorphism(s) of Figure 2 positioned below the polymorphic site in the sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the HRH2 gene. As described in more detail below, the inventors herein discovered 11 isogenes of the HRH2 gene by characterizing the HRH2 gene found in genomic DNAs isolated from an Index Repository that contains immortalized cell lines from one chimpanzee and 93 human individuals. The human individuals included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals) (CA), African descent (20 individuals) (AF), Asian (20 individuals) (AS), or Hispanic/Latino (17 individuals) (HL). To the extent possible, the members of this reference population were organized into population subgroups by the self-identified ethnogeographic origin of their four grandparents as shown in Table 1 below.

Table 1. Population Groups in the Index Repository

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

In addition, the Index Repository contains three unrelated indigenous American Indians (AM) (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family.

The HRH2 isogenes present in the human reference population are defined by haplotypes for 10 polymorphic sites in the HRH2 gene, all of which are believed to be novel. The HRH2

polymorphic sites identified by the inventors are referred to as PS1-PS10 to designate the order in which they are located in the gene (see Table 2 below). Using the genotypes identified in the Index Repository for PS1-PS10 and the methodology described in the Examples below, the inventors herein also determined the pair of haplotypes for the HRH2 gene present in individual human members of this repository. The human genotypes and haplotypes found in the repository for the HRH2 gene include those shown in Tables 3 and 4, respectively. The polymorphism and haplotype data disclosed herein are useful for validating whether HRH2 is a suitable target for drugs to treat acid-peptic disorders of the gastrointestinal tract, screening for such drugs and reducing bias in clinical trials of such drugs.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a

listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or

methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.

As discussed above, information on the identity of genotypes and haplotypes for the HRH2 gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is expected to be useful for a variety of drug discovery and development applications. Thus, the invention also provides compositions and methods for detecting the novel HRH2 polymorphisms and haplotypes identified herein.

The compositions comprise at least one HRH2 genotyping oligonucleotide. In one embodiment, a HRH2 genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the

art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of a HRH2 polynucleotide, i.e., a HRH2 isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-HRH2 polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the HRH2 gene using the polymorphism information provided herein in conjunction with the known sequence information for the HRH2 gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruaño et al., 87 *Proc. Natl. Acad. Sci. USA* 6296-6300, 1990. Typically, an ASO will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotides of the invention include ASO probes and ASO primers. ASO probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15mer, the 8th or 9th position in a 16mer, and the 10th or 11th position in a 20mer). An ASO primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. ASO probes and primers hybridizing to either the coding or noncoding strand are contemplated by the invention.

ASO probes and primers listed below use the appropriate nucleotide symbol (R= G or A, Y= T or C, M= A or C, K= G or T, S= G or C, and W= A or T; WIPO standard ST.25) at the position of the polymorphic site to represent the two alternative allelic variants observed at that polymorphic site.

A preferred ASO probe for detecting HRH2 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.: AB023486.1

CAGCTGCRTCTCCAC (SEQ ID NO:4) and its complement,
CTTCAGGRGACCGTC (SEQ ID NO:5) and its complement,
ACCGTCTRAGGACTG (SEQ ID NO:6) and its complement,
ACTCTACYGCATGCA (SEQ ID NO:7) and its complement,
CTGACCTRCTCCTCG (SEQ ID NO:8) and its complement,
TCAGCCTYGACCGGT (SEQ ID NO:9) and its complement,
ATGAAGTRTACGGGC (SEQ ID NO:10) and its complement,
GCCATCARTGAGGTG (SEQ ID NO:11) and its complement,
AGGATGGSGGCAATG (SEQ ID NO:12) and its complement, and
AGGAACTYTTCATGA (SEQ ID NO:13) and its complement.

A preferred ASO primer for detecting HRH2 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.:AB023486.1

GGGATACAGCTGCRT (SEQ ID NO:14); GGTCATGTGGAGAYG (SEQ ID NO:15);
CTAGCTCTTCAGGRG (SEQ ID NO:16); TCCTCAGACGGTCYC (SEQ ID NO:17);
CAGGGGACCGTCTRA (SEQ ID NO:18); AAACTCCAGTCCTYA (SEQ ID NO:19);
GCCTGGACTCTACYG (SEQ ID NO:20); TGATCTTGCATGCRG (SEQ ID NO:21);
CTATCACTGACCTRC (SEQ ID NO:22); GGAGGCCGAGGAGYA (SEQ ID NO:23);
TCATGATCAGCCTYG (SEQ ID NO:24); CGCAGTACCGGTCRA (SEQ ID NO:25);
AGGTCAATGAAGTRT (SEQ ID NO:26); CCACCAGCCCGTAYA (SEQ ID NO:27);
GATGATGCCATCART (SEQ ID NO:28); TTCTAACACCTCAYT (SEQ ID NO:29);
GTGCACAGGATGGSG (SEQ ID NO:30); CCCTCCCATTGCCSC (SEQ ID NO:31);
TGTTCTAGGAACTYT (SEQ ID NO:32); and AAGTGCTCATGAARA (SEQ ID NO:33).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such

oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3′-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site.

A particularly preferred oligonucleotide primer for detecting HRH2 gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.:AB023486.1

ATACAGCTGC (SEQ ID NO:34); CATGTGGAGA (SEQ ID NO:35);
GCTCTTCAGG (SEQ ID NO:36); TCAGACGGTC (SEQ ID NO:37);
GGGACCGTCT (SEQ ID NO:38); CTCCAGTCCT (SEQ ID NO:39);
TGGACTCTAC (SEQ ID NO:40); TCTTGCATGC (SEQ ID NO:41);
TCACTGACCT (SEQ ID NO:42); GGCCGAGGAG (SEQ ID NO:43);
TGATCAGCCT (SEQ ID NO:44); AGTACCGGTC (SEQ ID NO:45);
TCAATGAAGT (SEQ ID NO:46); CCAGCCCGTA (SEQ ID NO:47);
GATGCCATCA (SEQ ID NO:48); TAACACCTCA (SEQ ID NO:49);
CACAGGATGG (SEQ ID NO:50); TCCCATTGCC (SEQ ID NO:51);
TCTAGGAACT (SEQ ID NO:52); and TGCTCATGAA (SEQ ID NO:53).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

HRH2 genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized HRH2 genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for

genotyping and/or haplotyping the HRH2 gene in an individual. As used herein, the terms "HRH2 genotype" and "HRH2 haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the HRH2 gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid sample comprising the two copies of the HRH2 gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10 in the two copies to assign a HRH2 genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-PS10.

Typically, the nucleic acid sample is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen, saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid sample may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from a tissue in which the HRH2 gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ untranslated regions. If a HRH2 gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid sample containing only one of the two copies of the HRH2 gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10 in that copy to assign a HRH2 haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the HRH2 gene or fragment such as one of the methods described above for preparing HRH2 isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two HRH2 gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional HRH2 clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the HRH2 gene in an individual. In a particularly preferred embodiment, the nucleotide at each of PS1-PS10 is identified.

In another embodiment, the haplotyping method comprises determining whether an individual has one or more of the HRH2 haplotypes shown in Table 4. This can be accomplished by identifying,

for one or both copies of the individual's HRH2 gene, the phased sequence of nucleotides present at each of PS1-PS10. The present invention also contemplates that typically only a subset of PS1-PS10 will need to be directly examined to assign to an individual one or more of the haplotypes shown in Table 4. This is because at least one polymorphic site in a gene is frequently in strong linkage disequilibrium with one or more other polymorphic sites in that gene (Drysdale, CM et al. 2000 *PNAS* 97:10483-10488; Rieder MJ et al. 1999 *Nature Genetics* 22:59-62). Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stephens, JC 1999, *Mol. Diag.* 4:309-317). Techniques for determining whether any two polymorphic sites are in linkage disequilibrium are well-known in the art (Weir B.S. 1996 *Genetic Data Analysis II*, Sinauer Associates, Inc. Publishers, Sunderland, MA).

In a preferred embodiment, a HRH2 haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10 in each copy of the HRH2 gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-PS10 in each copy of the HRH2 gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the HRH2 gene, or a fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

The target region(s) may be amplified using any oligonucleotide-directed amplification

method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., *Proc. Natl. Acad. Sci. USA* 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., *Science* 241:1077-1080, 1988).

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* **89**:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the HRH2 gene of an individual may also be determined by hybridization of a nucleic acid sample containing one or both copies of the gene, or fragment(s) thereof, to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., *Proc. Natl. Acad. Sci. USA* 82:7575, 1985; Meyers et al., *Science* 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. *Ann. Rev. Genet.* 25:229-253,

1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., *Genomics* 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., *Nucl. Acids Res.* 18:2699-2706, 1990; Sheffield et al., *Proc. Natl. Acad. Sci. USA* 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., *Nucl. Acids Res.* 17:8392, 1989; Ruaño et al., *Nucl. Acids Res.* 19, 6877-6882, 1991; WO 93/22456; Turki et al., *J. Clin. Invest.* 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping another polymorphic site that is in linkage disequilibrium with the polymorphic site that is of interest. Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

In another aspect of the invention, an individual's HRH2 haplotype pair is predicted from its HRH2 genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying a HRH2 genotype for the individual at two or more HRH2 polymorphic sites described herein, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing HRH2 haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include the HRH2 haplotype pairs shown in Table 3.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see.

For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by 2n=log(1-q)/log(1-p) where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), single molecule dilution, or allele-specific long-range PCR (Michalotos-Beloin et al., *Nucleic Acids Res.* 24:4841-4843, 1996).

In one embodiment of this method for predicting a HRH2 haplotype pair for an individual, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. Alternatively, the haplotype pair in an individual may be predicted from the individual's genotype for that gene using reported methods (e.g., Clark et al. 1990 *Mol Bio Evol* 7:111-22) or through a commercial haplotyping service such as offered by Genaissance Pharmaceuticals, Inc. (New Haven, CT). In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER

System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., *supra*).

The invention also provides a method for determining the frequency of a HRH2 genotype, haplotype, or haplotype pair in a population. The method comprises, for each member of the population, determining the genotype or the haplotype pair for the novel HRH2 polymorphic sites described herein, and calculating the frequency any particular genotype, haplotype, or haplotype pair is found in the population. The population may be a reference population, a family population, a same sex population, a population group, or a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for HRH2 genotypes, haplotypes, and/or haplotype pairs are determined in a reference population and used in a method for identifying an association between a trait and a HRH2 genotype, haplotype, or haplotype pair. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s), haplotype(s), or haplotype pair(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s), haplotype(s), or haplotype pair(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes, haplotypes, and/or haplotype pairs observed in the populations are compared. If a particular HRH2 genotype, haplotype, or haplotype pair is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that HRH2 genotype, haplotype, or haplotype pair. Preferably, the HRH2 genotype, haplotype, or haplotype pair being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 3 and 4, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting HRH2 or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side

effects).

In order to deduce a correlation between clinical response to a treatment and a HRH2 genotype, haplotype, or haplotype pair, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the HRH2 gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and HRH2 genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their HRH2 genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the HRH2 gene give the most significant contribution to the

differences in phenotype. One regression model useful in the invention is described in PCT Application Serial No. PCT/US00/17540, entitled "Methods for Obtaining and Using Haplotype Data".

A second method for finding correlations between HRH2 haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., *supra,* Ch. 10), or other global or local optimization approaches (see discussion in Judson, *supra*) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in PCT Application Serial No. PCT/US00/17540.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the HRH2 gene. As described in PCT Application Serial No. PCT/US00/17540, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, *supra*, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of HRH2 genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the HRH2 gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the HRH2 gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying HRH2 genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

In another embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the HRH2 gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant HRH2 gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel

polymorphic sites PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10. Similarly, the nucleotide sequence of a variant fragment of the HRH2 gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence of the HRH2 gene, which is defined by haplotype 1, (or other reported HRH2 sequences) or to portions of the reference sequence (or other reported HRH2 sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of adenine at PS1, adenine at PS2, adenine at PS3, thymine at PS4, adenine at PS5, thymine at PS6, adenine at PS7, guanine at PS8, cytosine at PS9 and thymine at PS10. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the HRH2 gene which is defined by any one of haplotypes 2-11 shown in Table 4 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the HRH2 gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the novel polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

HRH2 isogenes may be isolated using any method that allows separation of the two "copies" of the HRH2 gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., *Proc. Natl. Acad. Sci.* 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 1989, *supra*; Ruaño et al., 1991, *supra*; Michalatos-Beloin et al., *supra*).

The invention also provides HRH2 genome anthologies, which are collections of HRH2 isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. A HRH2 genome anthology may comprise individual HRH2 isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the HRH2 isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred HRH2 genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 4 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded HRH2 protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant HRH2 sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 *Science* 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the HRH2 gene will produce HRH2 mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of a HRH2 cDNA comprising a nucleotide sequence which is a polymorphic variant of the HRH2 reference coding sequence shown in Figure 2. Thus, the invention also provides HRH2 mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO: 2 (Fig. 2), or its corresponding RNA sequence, except for having one or more polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 45, adenine at a position corresponding to nucleotide 195, thymine at a position corresponding to nucleotide 342, adenine at a position corresponding to nucleotide 543 and guanine at a position corresponding to nucleotide 797.

A particularly preferred polymorphic cDNA variant comprises the coding sequence of a HRH2 isogene defined by haplotypes 2-11. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized HRH2 cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

As used herein, a polymorphic variant of a HRH2 gene fragment comprises at least one novel polymorphism identified herein and has a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, such fragments are between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the HRH2 polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the HRH2 gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the HRH2 genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular HRH2 protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the HRH2 isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular HRH2 isogene. Expression of a HRH2 isogene may be turned off by transforming a targeted organ, tissue or cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of HRH2 mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze

the specific cleavage of HRH2 mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference HRH2 amino acid sequence shown in Figure 3. The location of a variant amino acid in a HRH2 polypeptide or fragment of the invention is identified by aligning its sequence against SEQ ID NO:3 (Fig.3). A HRH2 protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO:3 except for having serine at a position corresponding to amino acid position 266. The invention specifically excludes amino acid sequences identical to those previously identified for HRH2, including SEQ ID NO:3, and previously described fragments thereof. HRH2 protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO:3 and having serine at a position corresponding to amino acid position 266.

The invention also includes HRH2 peptide variants, which are any fragments of a HRH2 protein variant that contain serine at a position corresponding to amino acid position 266. A HRH2 peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such HRH2 peptide variants may be useful as antigens to generate antibodies specific for one of the above HRH2 isoforms. In addition, the HRH2 peptide variants may be useful in drug screening assays.

A HRH2 variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant HRH2 genomic and cDNA sequences as described above. Alternatively, the HRH2 protein variant may be isolated from a biological sample of an individual having a HRH2 isogene which encodes the variant protein. Where the sample contains two different HRH2 isoforms (i.e., the individual has different HRH2 isogenes), a particular HRH2 isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular HRH2 isoform but does not bind to the other HRH2 isoform.

The expressed or isolated HRH2 protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the HRH2 protein as discussed further below. HRH2 variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing,

gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant HRH2 gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric HRH2 protein. The non-HRH2 portion of the chimeric protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the HRH2 and non-HRH2 portions so that the HRH2 protein may be cleaved and purified away from the non-HRH2 portion.

An additional embodiment of the invention relates to using a novel HRH2 protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known HRH2 protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The HRH2 protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to a HRH2 variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the HRH2 protein(s) of interest and then washed. Bound HRH2 protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel HRH2 protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the HRH2 protein.

In yet another embodiment, when a particular HRH2 haplotype or group of HRH2 haplotypes encodes a HRH2 protein variant with an amino acid sequence distinct from that of HRH2 protein isoforms encoded by other HRH2 haplotypes, then detection of that particular HRH2 haplotype or group of HRH2 haplotypes may be accomplished by detecting expression of the encoded HRH2 protein variant using any of the methods described herein or otherwise commonly known to the skilled artisan.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel HRH2 variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The HRH2 protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the HRH2 protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel protein isoforms described herein is administered to an individual to neutralize activity of the HRH2 isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel protein isoforms described herein may be used to immunoprecipitate the HRH2 protein variant from solution as well as react with HRH2 protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect HRH2 protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel HRH2 protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the HRH2 protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in E. coli is the

subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci.USA 86;10029).

Effect(s) of the polymorphisms identified herein on expression of HRH2 may be investigated by preparing recombinant cells and/or nonhuman recombinant organisms, preferably recombinant animals, containing a polymorphic variant of the HRH2 gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into HRH2 protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired HRH2 isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the HRH2 isogene is introduced into a cell in such a way that it recombines with the endogenous HRH2 gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired HRH2 gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the HRH2 isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the HRH2 isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant nonhuman organisms, i.e., transgenic animals, expressing a variant HRH2 gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the HRH2 isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In:

Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human HRH2 isogene and producing human HRH2 protein can be used as biological models for studying diseases related to abnormal HRH2 expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel HRH2 isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel HRH2 isogenes; an antisense oligonucleotide directed against one of the novel HRH2 isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel HRH2 isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel HRH2 isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the HRH2 gene and its genomic

variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The HRH2 polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD-ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

EXAMPLE 1

This example illustrates examination of various regions of the HRH2 gene for polymorphic sites.

Amplification of Target Regions

The following target regions of the HRH2 gene were amplified using PCR primer pairs. The primers used for each region are represented below by providing the nucleotide positions of their initial and final nucleotides, which correspond to positions in the indicated GenBank Accession Number.

PCR Primer Pairs

GenBank Acc No.	Fragment No.	Forward Primer	Reverse Primer (complement of)	PCR Product
AB023486.1	Fragment 1	11746-11769	12233-12213	488 nt
	Fragment 2	12124-12146	12814-12794	691 nt
	Fragment 3	12566-12591	13124-13102	559 nt
	Fragment 4	12806-12829	13358-13336	553 nt

Fragment 5	13123-13142	13569-13546	447 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

Reaction volume	= 10 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 1 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	= 6.6µl

Amplification profile:
97°C - 2 min. 1 cycle

97°C - 15 sec.
70°C - 45 sec.
72°C - 45 sec. } 10 cycles

97°C - 15 sec.
64°C - 45 sec.
72°C - 45 sec. } 35 cycles

Sequencing of PCR Products

The PCR products were purified using a Whatman/Polyfiltronics 100 µl 384 well unifilter plate essentially according to the manufacturers protocol. The purified DNA was eluted in 50 µl of distilled water. Sequencing reactions were set up using Applied Biosystems Big Dye Terminator chemistry essentially according to the manufacturers protocol. The purified PCR products were sequenced in both directions using the primer sets described previously or those represented below by the nucleotide positions of their initial and final nucleotides, which correspond to positions in the indicated GenBank Accession Number. Reaction products were purified by isopropanol precipitation, and run on an Applied Biosystems 3700 DNA Analyzer.

Sequencing Primer Pairs

GenBank Acc No.	Fragment No.	Forward Primer	Reverse Primer (complement of)
AB023486.1	Fragment 1	11756-11775	12156-12139
	Fragment 2	12187-12206	12788-12770
	Fragment 3	12609-12626	12996-12977
	Fragment 4	12817-12836	13354-13337
	Fragment 5	13154 13172	13543-13526

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., *Nucleic Acids Res.* 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the HRH2 gene are listed in Table 2 below.

Table 2. Polymorphic Sites Identified in the HRH2 Gene

Polymorphic Site Number	PolyId[a]	Nucleotide Position	Reference Allele	Variant Allele
PS1	3485	11881(Acc#AB023486.1)	G	A
PS2	3486	12001(Acc#AB023486.1)	G	A
PS3	3487	12010(Acc#AB023486.1)	G	A
PS4	3494	12352(Acc#AB023486.1)	C	T
PS5	3493	12502(Acc#AB023486.1)	G	A
PS6	3492	12649(Acc#AB023486.1)	C	T
PS7	3495	12850(Acc#AB023486.1)	G	A
PS8	3491	13104(Acc#AB023486.1)	A	G
PS9	3489	13416(Acc#AB023486.1)	G	C
PS10	3488	13497(Acc#AB023486.1)	C	T

[a]PolyId is a unique identifier assigned to each PS by Genaissance Pharmaceuticals, Inc.

EXAMPLE 2

This example illustrates analysis of the HRH2 polymorphisms identified in the Index Repository for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 3 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 3, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 3 were inferred based on linkage disequilibrium and/or Mendelian inheritance.

Table 3. Genotypes and Haplotype Pairs Observed in the HRH2 Gene

Genotype	Polymorphic Sites										HAP	
Number	PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS9	PS10	PAIR	
1	G	G	G	C	G	C	G	A	G	C	1	1
2	G	G	G	C/T	G	C	G	A	G	C	1	2
3	G	G	G	C	G	C	G/A	A	G	C	1	3
4	G	G	G	C	G/A	C	G	A	G	C	1	5
5	G	G	G	C	G	C	G	A/G	G	C	1	6
6	G	G	G	C	G	C	G	A	G	C/T	1	7
7	G/A	G	G	C	G	C	G	A	G	C	1	8
8	G	G/A	G	C	G	C	G	A	G	C	1	9
9	G	G	G	C	-	-	-	-	G/C	C	1	10
10	G	G	G/A	-	-	-	-	-	-	-	1	11
11	G	G	G	T	G	C/T	G	A	G	C	2	4

The haplotype pairs shown in Table 3 were estimated from the unphased genotypes using a computer-implemented extension of Clark's algorithm (Clark, A.G. 1990 *Mol Bio Evol* 7, 111-122) for

assigning haplotypes to unrelated individuals in a population sample. In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is augmented with haplotypes obtained from two families (one three-generation Caucasian family and one two-generation African-American family) and then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 11 human HRH2 haplotypes shown in Table 4 below.

Table 4. Haplotypes Observed in the HRH2 Gene

Haplotype	Polymorphic Sites									
Number	PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS9	PS10
1	G	G	G	C	G	C	G	A	G	C
2	G	G	G	T	G	C	G	A	G	C
3	G	G	G	C	G	C	A	A	G	C
4	G	G	G	T	G	T	G	A	G	C
5	G	G	G	C	A	C	G	A	G	C
6	G	G	G	C	G	C	G	G	G	C
7	G	G	G	C	G	C	G	A	G	T
8	A	G	G	C	G	C	G	A	G	C
9	G	A	G	C	G	C	G	A	G	C
10	G	G	G	C	G	C	G	A	C	C
11	G	G	A	C	G	C	G	A	G	C

Table 5 below shows the number of chromosomes seen for each haplotype arranged by the ethnic background of the subjects in the Index Repository.

Table 5. Frequencies of Haplotypes Observed in the HRH2 Gene

Haplotype Number	AF	AS	CA	HL	AM	Total
1	32	36	39	27	6	140
2	4	0	0	2	0	6
3	0	1	1	2	0	4
4	1	0	0	0	0	1
5	0	0	1	0	0	1
6	0	0	1	0	0	1
7	1	0	0	0	0	1
8	0	0	0	1	0	1
9	1	0	0	0	0	1
10	1	0	0	0	0	1
11	0	1	0	0	0	1

Table 6 below shows the number of subjects assigned to each haplotype pair arranged by the ethnic background of the subjects in the Index Repository.

Table 6. Frequencies of Haplotype Pairs Observed in the HRH2 Gene							
HAP Pair		AF	AS	CA	HL	AM	Total
1	1	13	17	18	11	3	62
2	1	4	0	0	2	0	6
3	1	0	1	1	2	0	4
4	2	1	0	0	0	0	1
5	1	0	0	1	0	0	1
6	1	0	0	1	0	0	1
7	1	1	0	0	0	0	1
8	1	0	0	0	1	0	1
9	1	1	0	0	0	0	1
10	1	1	0	0	0	0	1
11	1	0	1	0	0	0	1

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. A method for haplotyping the Histamine H2 receptor (HRH2) gene of an individual which comprises determining whether the individual has one of the HRH2 haplotypes shown in Table 4 or one of the haplotype pairs shown in Table 3.

2. The method of claim 1, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS10 on at least one copy of the individual's HRH2 gene.

3. The method of claim 1, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS1-PS10 on both copies of the individual's HRH2 gene.

4. A method for genotyping the Histamine H2 receptor (HRH2) gene of an individual, comprising determining for the two copies of the HRH2 gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10.

5. The method of claim 4, wherein the determining step comprises:
   (a) isolating from the individual a nucleic acid mixture comprising both copies of the HRH2 gene, or a fragment thereof, that are present in the individual;
   (b) amplifying from the nucleic acid mixture a target region containing the selected polymorphic site;
   (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
   (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the selected polymorphic site; and
   (e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

6. The method of claim 4, which comprises determining for the two copies of the HRH2 gene present in the individual the identity of the nucleotide pair at each of PS1-PS10.

7. A method for haplotyping the Histamine H2 receptor (HRH2) gene of an individual which comprises determining, for one copy of the HRH2 gene present in the individual, the identity of the nucleotide at two or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10.

8. The method of claim 7, wherein the determining step comprises:
   (a) isolating from the individual a nucleic acid sample containing only one of the two copies of the HRH2 gene, or a fragment thereof, that is present in the individual;
   (b) amplifying from the nucleic acid molecule a target region containing the selected

polymorphic site;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the selected polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

9. A method for predicting a haplotype pair for the Histamine H2 receptor (HRH2) gene of an individual comprising:

(a) identifying a HRH2 genotype for the individual, wherein the genotype comprises the nucleotide pair at two or more polymorphic sites selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) comparing the possible haplotype pairs to the data in Table 3; and

(d) assigning a haplotype pair to the individual that is consistent with the data.

10. The method of claim 9, wherein the identified genotype of the individual comprises the nucleotide pair at each of PS1-PS10.

11. A method for identifying an association between a trait and at least one haplotype or haplotype pair of the Histamine H2 receptor (HRH2) gene which comprises comparing the frequency of the haplotype or haplotype pair in a population exhibiting the trait with the frequency of the haplotype or haplotype pair in a reference population, wherein the haplotype is selected from haplotypes 1-11 shown in Table 4 and the haplotype pair is selected from the haplotype pairs shown in Table 3, wherein a higher frequency of the haplotype or haplotype pair in the trait population than in the reference population indicates the trait is associated with the haplotype or haplotype pair.

12. The method of claim 11, wherein the trait is a clinical response to a drug targeting HRH2.

13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the Histamine H2 receptor (HRH2) gene at a polymorphic site selected from the group consisting of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10.

14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the HRH2 gene at a region containing the polymorphic site.

15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS:4-13, the complements of SEQ ID NOS:4-13, and SEQ ID NOS:14-33.

16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension

oligonucleotide.

17. The composition of claim 16, wherein the primer extension oligonucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NOS:34-53.

18. A kit for genotyping the HRH2 gene of an individual, which comprises a set of oligonucleotides designed to genotype each of PS1, PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9 and PS10.

19. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for the Histamine H2 receptor (HRH2) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises a HRH2 isogene defined by a haplotype selected from the group consisting of haplotypes 1-11 in Table 4; and

(b) a second nucleotide sequence which is complementary to the first nucleotide sequence.

20. The isolated polynucleotide of claim 19, which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

21. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 19, wherein the organism expresses a HRH2 protein encoded by the first nucleotide sequence.

22. The recombinant organism of claim 21, which is a nonhuman transgenic animal.

23. The isolated polynucleotide of claim 19, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the HRH2 gene, the fragment comprising one or more polymorphisms selected from the group consisting of adenine at PS1, adenine at PS2, adenine at PS3, thymine at PS4, adenine at PS5, thymine at PS6, adenine at PS7, guanine at PS8, cytosine at PS9 and thymine at PS10.

24. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the HRH2 cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises the coding sequence of a HRH2 isogene defined by one of the haplotypes shown in Table 4.

25. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 24, wherein the organism expresses a Histamine H2 receptor (HRH2) protein encoded by the polymorphic variant sequence.

26. The recombinant organism of claim 25, which is a nonhuman transgenic animal.

27. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the HRH2 protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:3 and the polymorphic variant is encoded by an isogene defined by one

of the haplotypes shown in Table 4.

28. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 27.

29. A method for screening for drugs targeting the isolated polypeptide of claim 27 which comprises contacting the HRH2 polymorphic variant with a candidate agent and assaying for binding activity.

30. A computer system for storing and analyzing polymorphism data for the Histamine H2 receptor gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 3 and the haplotypes shown in Table 4.

31. A genome anthology for the Histamine H2 receptor (HRH2) gene which comprises HRH2 isogenes defined by any one of haplotypes 1-11 shown in Table 4.

POLYMORPHISMS IN THE HRH2 GENE (Accession No. AB023486.1)

AAGGACTCCA CTGAGAAGAG AGTGCTCCCA GGGCCTTGAA AACTCAAAAC
TCTGCCCAAT GGGATTAAAA AAACACCCCC TTTCTGTCCA CCTCAGAAAG 8400
CACAAGCCTC TGCCTGAGGT GAGTACAAGG GAGGTAGAAG TGAGAGAGGA
TGGAGAACCT ACCTCTGGAA GACTTCCTGG AGGAGGTGAC AGAGGAGGCC 8500
AGGGTTTGAA AAAAGAGTAA GAGCTCTCTA GGTAAGGAAA GACCTTCCAG
GCAGTGAAAA CAGCATGTGC AAAGGCACAG AATCAGAATG GGCCTGGCAC 8600
TTTTCAGAAT TAAATGCCGC AGAGCCAGAT CAACACTCCA AGGCAGCAGG
TGCTGTACAA TCTAACTCGG TACAAAGTGC AGAGAGATTA CTCCTGCTTT 8700
TAGGCATCTA GGAGGACTCA AGGAGACAGT TACCTGAGCT GGCCTTGAAA
GGAGAGTGGT TGAAGTGAAT ACAAGAAAGA GTGACACACC AGGCATAGGG 8800
GATGCAAAGG CAGAGGAAGG AGGGAGAAGG TGTGCTGCAT CTGGGGACAT
GCTCGGTTGC CTGAGTGTCT GTGAAATTAA CAGGCAGATG CTGGAGTGCT 8900
GGATGAGACC CTGTTGGAAA GACAGGTTGA AGTCTGATCA AAGAACGGGT
ATTAAGTACA GGCCAAAGGA GTTTGAGCTG AATTTGAATC CAGGTTCTGC 9000
CACTTTCTAG GGGTGTGACA TCACCTGGCC TCTTTGAGCC TCAGTCTCCT
CAACCGTTAA ATGGAGACAT GATTTTCCTC ACTCGTGGGG TCGTGAGGAT 9100
TCAGTGATAT TTTGTTTACT ATGCATGGTA TCGCTTTGCC CAGGAGGATC
CAGCAGGACA GTTGTGCCGT GGTTCATTAA ACCTTCTTCC CCCTGCAGTT 9200
GGCAGCTTAG GTTGTTTTCC ATTCTTGCCA TTGCACGTCT GCCAACGGAC
ATCCTTGTGC ACATATCCTC ATGAGAGCAG GTGCATTGCC ACGGGCTGGG 9300
CTCCCGAAAG AATGGCTTCG ATCTAAGTGA CCTAGAGGGC CATCCAGTAC
ATATTGGGAA TTTTACAAGA AGTGAAGGAG AGGCAGTTGC TGAGTAGCAT 9400
GTATGCTGTG TGCGGGCAGT GCCCAGTGTG TGCCCATTGT GCATGCAGAG
GCATGAGAAA CCCAGGAGGC CCCCACATAC CAAACAGTCA GCAGCAGTCA 9500
CTTCCTGGTG GAGGGGCTCA GGTGTGTGGT GGCTCTGATT TTTGGCCCTT
TCTGTACACC TCAGGGATAT TTGGCTTACT GGAATAAGCC TGTCATCACT 9600
TTGGAAGTTT GAAAGGACAG TTTGAGAGAC TATTTTTTGT TAGCTATGAT
GGTGACAATG ACGGTGACAA GGGTGACATG ATTCTTCCTC ACAGGCAAAG 9700
AGAACCCATT GAAGGTTGCT CCATAAGGGG ATGACGTGGC CCTCTAGGGC
AATCCTTCTG CCCACATGGG ACAGGAGAGG ATGGGAGGGA GAGAGAGCGC 9800
CGTACAAAAT CCCCTGCCCT CGAAAAAGGT GACCAGAACT TACGGGAGCT
GGGCTTTAAA TAGTAATACA GCAGAAGTGG TGCTTCTTTG AAAACTCTAC 9900
GAAGAAAGGA CTCACAGCTG TTTTGGTCAT CATTGTCCCC CCAGCACCTG
GAGTAGCATC TGACACATTG TGGGTTTTCA AAGAAATATG TGTTGAAAAG 10000
GACTGAGGAA GGGAAGGAGG GGCGGGAGAG GAGAGAAGGA AAGGATTTTC
TAATTATAAG AGAGGCTGGT CCACTCAGAT TTGAAGACCC CTCATCCAAA 10100
CATGAACTCA AATGTCAGCC TGATGCCAAA ATACCTATAC CCGGCTGTTG
TGGTAACTCA GAAAAACTTT GATGAGAAAT AGGGATTCTC CAAAAAAATA 10200
CCCTTATTGC TACCCCCTCC AAAGTTAATT AAGTGTGGCG GTGAGACCCA
TGCTGTTTGG GCTATTTATG TAATGAAAAG CTCTGCTAAA AATAGTACCC 10300
GGAGGTTCAG CTGCCAGGAC TGGTGGCTGA GCCAGGGAGC GTGGCAGAAG
GTGATCGGGA ACCCTGATCC CTGCACGAAC ACAATCGTTG TGATTAAATG 10400
ATGCTTCTCA TATTCAGGTT GCAAACCGCT CCAGGCAAGA GCCTGGGAGA
GGCTTAATTT ATGTTTTATG GAGATTAGCT TGCTTCTGAG AGGGGGTTCC 10500
GCCATCAACT CGATTTCTGC ACACACTCAG AGCTCAGGAG ATTGCCTGCT
CTCTGAAAAA CCAGGAGGAG CTGGTAGGGA AGGACTGGGG CGTGGAACAG 10600
GGCCTCAGCT TGTTTCTCGC CAGTTGGAAT CATATACAGC CACCACTGGG
AGGGAGCCTT GGAGATTACC TAGTGACCAG GGGCTCTTAT TTAATCAAGA 10700
GGCATTGCAC ATGCAGTTAC TTTATTGACG TAGTGTTTGC ACTACTTCAA
TAGTGCAAAG GGAGGAAGGG AGGAAGGAGA AACCACCCTA ATGCCCGTGA 10800
GAGGGTGATG TGGGGATGAT AGGATAAACC TATTCATGGT GAGCCTCCCC

FIGURE 1A

```
ACAGACTGTT ATGCAACCAT TAGAAGGAAT GAGTGAGTTC TGTCCCTGCT      10900
GACCTGGATG TGGGATCACG TGTTCTGCCA AGTAAGACCA GTAAGATTTA
GAGATGCGTT CGGAGCAGGG GTGGCGTGGA TGACAGGTGA CGAACTGCCG      11000
ATTAAATCTG GCCCTCTGCC TGGCTTTGTA CAGCCCGTGG CTAAGAATGG
TTTTTACATT TTTAAATTTT GGAAAAAAAT CAAAAAAAGG AGATTTTTAT      11100
GACACGTGAA AATGATATGA AATTTACATT TCAGTGTTCA TTACTGAAGT
TTTGTTGGAG TGCAGCCACG CTCTTCTGTC GGCACGTCAT CTGCATAGCT      11200
GCATTCGCAC TGCAAAGGCA GAGCCGAGCC GTCACAGGCT GTGAGGCAGA
ACGTGTTTAT TACCTGACCC TTTTCTGAAA AAGTTTGTCG GCCACCGCCT      11300
TAGAGTATGA TGACATTTTG CCAAACAGCA GCACCAGATA TGGTTCTCAG
CACTTCAGAG GAGTAGATTC ATTACATTTG TTTATTCGTT CATTCATTTT      11400
ATCCTTTATT CGTCAGATGC TTACTGTCTA TCTGTGGGGC CAGGCACCAG
GAACCAGCAG AGAACAAACA GATGCAGGCC TTCTTCTTGT GGAGCTGAAA      11500
GTCTGTGTCT AGATCCTTCT TTTCTCAGTA CAAGCACTTC CAGCCTTTCC
AACTCTTTCT CACGAGTCTC AGTGTCCAGC CCCTTGCATC CTGCCTCCCT      11600
TCTGCCTCCC CAGTATGTAT CCTGATTGCC CAAGTCCCTC CCAAAAAGAG
GTATCCAAAC TGGTCCCATA CTTCCTGGGG TCTTGCCAAG AGGTTGAGAT      11700
CTAGTAGAAG GACACATCTT GAATTGGGTC ATGCTTTCTA TCTGGTTTCA
AGGTGCTTAA CATCCAACCT TTGCCTTTTC AGCTCCTGCC CTCCACTGAC      11800
TCCAGAGAGG GAGATCCCCA GTACTTGACT CCATCACGCA GATGGGAGCA
GGCACCAGCT ATGGAGAGGG ATACAGCTGC GTCTCCACAT GACCCATCCT      11900
                                  A
GCATGACACC AAAGCCACCG CCAGACAGTG CCTCGGATTC TATGCAAAAC
CTGGGAAGCG GAGACCTACC CCAGCCCCGG GAGGAAGCTA GCTCTTCAGG      12000
GGACCGTCTG AGGACTGGAG TTTGATCCAT GAACCTGGCT TCGAGGCCTT
A       A
GCTTTTCTCT CTTCTTCATT CATATTCATT CCCAACACCT TAGAAGGTGT      12100
TGCTTAATTT ATTTCTAGAA AAGCAGCCCA GAGTCAGTCA TTGAAGCCTT
CCCCACCCCC TGGCCAAAAA AAAAAAAAAA AAAAAACTGG ACACATTTTG      12200
GATCTGTTGG GAGCTTGGAG TCCAGTGGTT GGCATAGTTG TCACATTGGG
AGCAGAGAAG AAGCAACCAG GGGCCCTGAT CAGGGGACTG AGCCGTAGAG      12300
TCCCAGGATG GCACCCAATG GCACAGCCTC TTCCTTTTGC CTGGACTCTA
     [EXON 1: 12308..
CCGCATGCAA GATCACCATC ACCGTGGTCC TTGCGGTCCT CATCCTCATC      12400
 T
ACCGTTGCTG GCAATGTGGT CGTCTGTCTG GCCGTGGGCT TGAACCGCCG
GCTCCGCAAC CTGACCAATT GTTTCATCGT GTCCTTGGCT ATCACTGACC      12500
TGCTCCTCGG CCTCCTGGTG CTGCCCTTCT CTGCCATCTA CCAGCTGTCC
 A
TGCAAGTGGA GCTTTGGCAA GGTCTTCTGC AATATCTACA CCAGCCTGGA      12600
TGTGATGCTC TGCACAGCCT CCATTCTTAA CCTCTTCATG ATCAGCCTCG
                                                    T
ACCGGTACTG CGCTGTCATG GACCCACTGC GGTACCCTGT GCTGGTCACC      12700
CCAGTTCGGG TCGCCATCTC TCTGGTCTTA ATTTGGGTCA TCTCCATTAC
CCTGTCCTTT CTGTCTATCC ACCTGGGGTG GAACAGCAGG AACGAGACCA      12800
GCAAGGGCAA TCATACCACC TCTAAGTGCA AAGTCCAGGT CAATGAAGTG
                                                    A
TACGGGCTGG TGGATGGGCT GGTCACCTTC TACCTCCCGC TACTGATCAT      12900
GTGCATCACC TACTACCGCA TCTTCAAGGT CGCCCGGGAT CAGGCCAAGA
GGATCAATCA CATTAGCTCC TGGAAGGCAG CCACCATCAG GGAGCACAAA      13000
GCCACAGTGA CACTGGCCGC CGTCATGGGG GCCTTCATCA TCTGCTGGTT
TCCCTACTTC ACCGCGTTTG TGTACCGTGG GCTGAGAGGG GATGATGCCA      13100
```

FIGURE 1B

```
TCAATGAGGT GTTAGAAGCC ATCGTTCTGT GGCTGGGCTA TGCCAACTCA
   G
GCCCTGAACC CCATCCTGTA TGCTGCGCTG AACAGAGACT TCCGCACCGG      13200
GTACCAACAG CTCTTCTGCT GCAGGCTGGC CAACCGCAAC TCCCACAAAA
CTTCTCTGAG GTCCAACGCC TCTCAGCTGT CCAGGACCCA AAGCCGAGAA      13300
CCCAGGCAAC AGGAAGAGAA ACCCCTGAAG CTCCAGGTGT GGAGTGGGAC
AGAAGTCACG GCCCCCCAGG GAGCCACAGA CAGGTAATAG CCCTAGCCAT      13400
                ..13387]
TGGTGCACAG GATGGGGGCA ATGGGAGGGG ATGCTACTGA TGGGAATGAT
                C
TAAGGGAGCT GCTGTTTAGG TGGTGCTGGT TTATGTTCTA GGAACTCTTC      13500
                                                    T
ATGAGCACTT TGTAAACACC CTCTTGCTTA ATCCTCCCAA CGGCCCCCAA
AGGTAGAACT TAGCTCCCTT TTAAAAGGAG CACATTAAAA TTCTCAGAGG      13600
ACTTGGCAAG GGCCGCACAG CTGGGGCATG AATTC                      13635
```

FIGURE 1C

POLYMORPHISMS IN THE CODING SEQUENCE OF HRH2

```
ATGGCACCCA ATGGCACAGC CTCTTCCTTT TGCCTGGACT CTACCGCATG
                                                  T
CAAGATCACC ATCACCGTGG TCCTTGCGGT CCTCATCCTC ATCACCGTTG      100
CTGGCAATGT GGTCGTCTGT CTGGCCGTGG GCTTGAACCG CCGGCTCCGC
AACCTGACCA ATTGTTTCAT CGTGTCCTTG GCTATCACTG ACCTGCTCCT      200
                                                  A
CGGCCTCCTG GTGCTGCCCT TCTCTGCCAT CTACCAGCTG TCCTGCAAGT
GGAGCTTTGG CAAGGTCTTC TGCAATATCT ACACCAGCCT GGATGTGATG      300
CTCTGCACAG CCTCCATTCT TAACCTCTTC ATGATCAGCC TCGACCGGTA
                                                  T
CTGCGCTGTC ATGGACCCAC TGCGGTACCC TGTGCTGGTC ACCCCAGTTC      400
GGGTCGCCAT CTCTCTGGTC TTAATTTGGG TCATCTCCAT TACCCTGTCC
TTTCTGTCTA TCCACCTGGG GTGGAACAGC AGGAACGAGA CCAGCAAGGG      500
CAATCATACC ACCTCTAAGT GCAAAGTCCA GGTCAATGAA GTGTACGGGC
                                                  A
TGGTGGATGG GCTGGTCACC TTCTACCTCC CGCTACTGAT CATGTGCATC      600
ACCTACTACC GCATCTTCAA GGTCGCCCGG GATCAGGCCA AGAGGATCAA
TCACATTAGC TCCTGGAAGG CAGCCACCAT CAGGGAGCAC AAAGCCACAG      700
TGACACTGGC CGCCGTCATG GGGGCCTTCA TCATCTGCTG GTTTCCCTAC
TTCACCGCGT TTGTGTACCG TGGGCTGAGA GGGGATGATG CCATCAATGA      800
                                                  G
GGTGTTAGAA GCCATCGTTC TGTGGCTGGG CTATGCCAAC TCAGCCCTGA
ACCCCATCCT GTATGCTGCG CTGAACAGAG ACTTCCGCAC CGGGTACCAA      900
CAGCTCTTCT GCTGCAGGCT GGCCAACCGC AACTCCCACA AAACTTCTCT
GAGGTCCAAC GCCTCTCAGC TGTCCAGGAC CCAAAGCCGA GAACCCAGGC     1000
AACAGGAAGA GAAACCCCTG AAGCTCCAGG TGTGGAGTGG GACAGAAGTC
ACGGCCCCCC AGGGAGCCAC AGACAGGTAA                          1080
```

FIGURE 2

ISOFORMS OF THE HRH2 PROTEIN

```
MAPNGTASSF CLDSTACKIT ITVVLAVLIL ITVAGNVVVC LAVGLNRRLR
NLTNCFIVSL AITDLLLGLL VLPFSAIYQL SCKWSFGKVF CNIYTSLDVM      100
LCTASILNLF MISLDRYCAV MDPLRYPVLV TPVRVAISLV LIWVISITLS
FLSIHLGWNS RNETSKGNHT TSKCKVQVNE VYGLVDGLVT FYLPLLIMCI      200
TYYRIFKVAR DQAKRINHIS SWKAATIREH KATVTLAAVM GAFIICWFPY
FTAFVYRGLR GDDAINEVLE AIVLWLGYAN SALNPILYAA LNRDFRTGYQ      300
                  S
QLFCCRLANR NSHKTSLRSN ASQLSRTQSR EPRQQEEKPL KLQVWSGTEV
TAPQGATDR                                                   359
```

SEQUENCE LISTING

<110> Genaissance Pharmaceuticals, Inc.
Chew, Anne
Choi, Julie Y.
Koshy, Beena

<120> Haplotypes of the HRH2 Gene

<130> MWH-0516PCT HRH2

<140> TBA
<141> 2001-04-12

<150> 60/196,406
<151> 2000-04-12

<160> 54

<170> PatentIn Ver. 2.1

<210> 1
<211> 5335
<212> DNA
<213> Homo sapien

<400> 1
aaggactcca ctgagaagag agtgctccca gggccttgaa aactcaaaac tctgcccaat 60
gggattaaaa aaacaccccc tttctgtcca cctcagaaag cacaagcctc tgcctgaggt 120
gagtacaagg gaggtagaag tgagagagga tggagaacct acctctggaa gacttcctgg 180
aggaggtgac agaggaggcc agggtttgaa aaaagagtaa gagctctcta ggtaaggaaa 240
gaccttccag gcagtgaaaa cagcatgtgc aaaggcacag aatcagaatg ggcctggcac 300
ttttcagaat taaatgccgc agagccagat caacactcca aggcagcagg tgctgtacaa 360
tctaactcgg tacaaagtgc agagagatta ctcctgcttt taggcatcta ggaggactca 420
aggagacagt tacctgagct ggccttgaaa ggagagtggt tgaagtgaat acaagaaaga 480
gtgacacacc aggcataggg gatgcaaagg cagaggaagg agggagaagg tgtgctgcat 540
ctggggacat gctcggttgc ctgagtgtct gtgaaattaa caggcagatg ctggagtgct 600
ggatgagacc ctgttggaaa gacaggttga agtctgatca aagaacgggt attaagtaca 660
ggccaaagga gtttgagctg aatttgaatc caggttctgc cactttctag gggtgtgaca 720
tcacctggcc tctttgagcc tcagtctcct caaccgttaa atggagacat gattttcctc 780
actcgtgggg tcgtgaggat tcagtgatat tttgtttact atgcatggta tcgctttgcc 840
caggaggatc cagcaggaca gttgtgccgt ggttcattaa accttcttcc ccctgcagtt 900
ggcagcttag gttgttttcc attcttgcca ttgcacgtct gccaacggac atccttgtgc 960
acatatcctc atgagagcag gtgcattgcc acgggctggg ctcccgaaag aatggcttcg 1020
atctaagtga cctagagggc catccagtac atattgggaa ttttacaaga agtgaaggag 1080
aggcagttgc tgagtagcat gtatgctgtg tgcgggcagt gcccagtgtg tgcccattgt 1140
gcatgcagag gcatgagaaa cccaggaggc ccccacatac caaacagtca gcagcagtca 1200
cttcctggtg gaggggctca ggtgtgtggt ggctctgatt tttggccctt tctgtacacc 1260

tcagggatat ttggcttact ggaataagcc tgtcatcact ttggaagttt gaaaggacag 1320
tttgagagac tatttttttgt tagctatgat ggtgacaatg acggtgacaa gggtgacatg 1380
attcttcctc acaggcaaag agaacccatt gaaggttgct ccataagggg atgacgtggc 1440
cctctagggc aatccttctg cccacatggg acaggagagg atgggaggga gagagagcgc 1500
cgtacaaaat cccctgccct cgaaaaaggt gaccagaact tacgggagct gggctttaaa 1560
tagtaataca gcagaagtgg tgcttctttg aaaactctac gaagaaagga ctcacagctg 1620
ttttggtcat cattgtcccc ccagcacctg gagtagcatc tgacacattg tgggttttca 1680
aagaaatatg tgttgaaaag gactgaggaa gggaaggagg ggcgggagag gagagaagga 1740
aaggattttc taattataag agaggctggt ccactcagat ttgaagaccc ctcatccaaa 1800
catgaactca aatgtcagcc tgatgccaaa atacctatac ccggctgttg tggtaactca 1860
gaaaaacttt gatgagaaat agggattctc caaaaaaata cccttattgc taccccctcc 1920
aaagttaatt aagtgtggcg gtgagaccca tgctgttttgg gctatttatg taatgaaaag 1980
ctctgctaaa aatagtaccc ggaggttcag ctgccaggac tggtggctga gccagggagc 2040
gtggcagaag gtgatcggga accctgatcc ctgcacgaac acaatcgttg tgattaaatg 2100
atgcttctca tattcaggtt gcaaaccgct ccaggcaaga gcctgggaga ggcttaattt 2160
atgttttatg gagattagct tgcttctgag agggggttcc gccatcaact cgatttctgc 2220
acacactcag agctcaggag attgcctgct ctctgaaaaa ccaggaggag ctggtaggga 2280
aggactgggg cgtggaacag ggcctcagct tgtttctcgc cagttggaat catatacagc 2340
caccactggg agggagcctt ggagattacc tagtgaccag gggctcttat ttaatcaaga 2400
ggcattgcac atgcagttac tttattgacg tagtgtttgc actacttcaa tagtgcaaag 2460
ggaggaaggg aggaaggaga aaccacccta atgcccgtga gagggtgatg tggggatgat 2520
aggataaacc tattcatggt gagcctcccc acagactgtt atgcaaccat tagaaggaat 2580
gagtgagttc tgtccctgct gacctggatg tgggatcacg tgttctgcca agtaagacca 2640
gtaagattta gagatgcgtt cggagcaggg gtggcgtgga tgacaggtga cgaactgccg 2700
attaaatctg gccctctgcc tggctttgta cagcccgtgg ctaagaatgg tttttacatt 2760
tttaaatttt ggaaaaaaat caaaaaaagg agatttttat gacacgtgaa aatgatatga 2820
aatttacatt tcagtgttca ttactgaagt tttgttggag tgcagccacg ctcttctgtc 2880
ggcacgtcat ctgcatagct gcattcgcac tgcaaaggca gagccgagcc gtcacaggct 2940
gtgaggcaga acgtgtttat tacctgaccc ttttctgaaa aagtttgtcg gccaccgcct 3000
tagagtatga tgacattttg ccaaacagca gcaccagata tggttctcag cacttcagag 3060
gagtagattc attacatttg tttattcgtt cattcatttt atcctttatt cgtcagatgc 3120
ttactgtcta tctgtggggc caggcaccag gaaccagcag agaacaaaca gatgcaggcc 3180
ttcttcttgt ggagctgaaa gtctgtgtct agatccttct tttctcagta caagcacttc 3240
cagcctttcc aactctttct cacgagtctc agtgtccagc cccttgcatc ctgcctcccc 3300
tctgcctccc cagtatgtat cctgattgcc caagtccctc ccaaaaagag gtatccaaac 3360
tggtcccata cttcctgggg tcttgccaag aggttgagat ctagtagaag gacacatctt 3420
gaattgggtc atgctttcta tctggtttca aggtgcttaa catccaacct ttgccttttc 3480
agctcctgcc ctccactgac tccagagagg gagatcccca gtacttgact ccatcacgca 3540
gatgggagca ggcaccagct atggagaggg atacagctgc gtctccacat gacccatcct 3600
gcatgacacc aaagccaccg ccagacagtg cctcggattc tatgcaaaac ctgggaagcg 3660
gagacctacc ccagccccgg gaggaagcta gctcttcagg ggaccgtctg aggactggag 3720
tttgatccat gaacctggct tcgaggcctt gcttttctct cttcttcatt catattcatt 3780
cccaacacct tagaaggtgt tgcttaattt atttctagaa aagcagccca gagtcagtca 3840
ttgaagcctt ccccaccccc tggccaaaaa aaaaaaaaaa aaaaaactgg acacattttg 3900
gatctgttgg gagcttggag tccagtggtt ggcatagttg tcacattggg agcagagaag 3960
aagcaaccag gggccctgat caggggactg agccgtagag tcccaggatg gcacccaatg 4020
gcacagcctc ttccttttgc ctggactcta ccgcatgcaa gatcaccatc accgtggtcc 4080
ttgcggtcct catcctcatc accgttgctg gcaatgtggt cgtctgtctg gccgtgggct 4140

tgaaccgccg gctccgcaac ctgaccaatt gtttcatcgt gtccttggct atcactgacc 4200
tgctcctcgg cctcctggtg ctgcccttct ctgccatcta ccagctgtcc tgcaagtgga 4260
gctttggcaa ggtcttctgc aatatctaca ccagcctgga tgtgatgctc tgcacagcct 4320
ccattcttaa cctcttcatg atcagcctcg accggtactg cgctgtcatg gacccactgc 4380
ggtaccctgt gctggtcacc ccagttcggg tcgccatctc tctggtctta atttgggtca 4440
tctccattac cctgtccttt ctgtctatcc acctggggtg gaacagcagg aacgagacca 4500
gcaagggcaa tcataccacc tctaagtgca aagtccaggt caatgaagtg tacgggctgg 4560
tggatgggct ggtcaccttc tacctcccgc tactgatcat gtgcatcacc tactaccgca 4620
tcttcaaggt cgcccgggat caggccaaga ggatcaatca cattagctcc tggaaggcag 4680
ccaccatcag ggagcacaaa gccacagtga cactggccgc cgtcatgggg gccttcatca 4740
tctgctggtt tccctacttc accgcgtttg tgtaccgtgg gctgagaggg gatgatgcca 4800
tcaatgaggt gttagaagcc atcgttctgt ggctgggcta tgccaactca gccctgaacc 4860
ccatcctgta tgctgcgctg aacagagact tccgcaccgg gtaccaacag ctcttctgct 4920
gcaggctggc caaccgcaac tcccacaaaa cttctctgag gtccaacgcc tctcagctgt 4980
ccaggaccca aagccgagaa cccaggcaac aggaagagaa acccctgaag ctccaggtgt 5040
ggagtgggac agaagtcacg gccccccagg gagccacaga caggtaatag ccctagccat 5100
tggtgcacag gatgggggca atgggagggg atgctactga tgggaatgat taagggagct 5160
gctgtttagg tggtgctggt ttatgttcta ggaactcttc atgagcactt tgtaaacacc 5220
ctcttgctta atcctcccaa cggcccccaa aggtagaact tagctccctt ttaaaaggag 5280
cacattaaaa ttctcagagg acttggcaag ggccgcacag ctggggcatg aattc 5335

<210> 2
<211> 1080
<212> DNA
<213> Homo sapien

<400> 2
atggcaccca atggcacagc ctcttccttt tgcctggact ctaccgcatg caagatcacc 60
atcaccgtgg tccttgcggt cctcatcctc atcaccgttg ctggcaatgt ggtcgtctgt 120
ctggccgtgg gcttgaaccg ccggctccgc aacctgacca attgtttcat cgtgtccttg 180
gctatcactg acctgctcct cggcctcctg gtgctgccct tctctgccat ctaccagctg 240
tcctgcaagt ggagctttgg caaggtcttc tgcaatatct acaccagcct ggatgtgatg 300
ctctgcacag cctccattct taacctcttc atgatcagcc tcgaccggta ctgcgctgtc 360
atggacccac tgcggtaccc tgtgctggtc accccagttc gggtcgccat ctctctggtc 420
ttaatttggg tcatctccat taccctgtcc tttctgtcta tccacctggg gtggaacagc 480
aggaacgaga ccagcaaggg caatcatacc acctctaagt gcaaagtcca ggtcaatgaa 540
gtgtacgggc tggtggatgg gctggtcacc ttctacctcc cgctactgat catgtgcatc 600
acctactacc gcatcttcaa ggtcgcccgg gatcaggcca agaggatcaa tcacattagc 660
tcctggaagg cagccaccat cagggagcac aaagccacag tgacactggc cgccgtcatg 720
ggggccttca tcatctgctg gtttccctac ttcaccgcgt ttgtgtaccg tgggctgaga 780
ggggatgatg ccatcaatga ggtgttagaa gccatcgttc tgtggctggg ctatgccaac 840
tcagccctga accccatcct gtatgctgcg ctgaacagag acttccgcac cgggtaccaa 900
cagctcttct gctgcaggct ggccaaccgc aactcccaca aaacttctct gaggtccaac 960
gcctctcagc tgtccaggac ccaaagccga gaacccaggc aacaggaaga gaaacccctg 1020
aagctccagg tgtggagtgg gacagaagtc acggcccccc agggagccac agacaggtaa 1080

<210> 3
<211> 359
<212> PRT
<213> Homo sapien

<400> 3

```
Met Ala Pro Asn Gly Thr Ala Ser Ser Phe Cys Leu Asp Ser Thr Ala
  1               5                  10                  15

Cys Lys Ile Thr Ile Thr Val Val Leu Ala Val Leu Ile Leu Ile Thr
            20                  25                  30

Val Ala Gly Asn Val Val Val Cys Leu Ala Val Gly Leu Asn Arg Arg
        35                  40                  45

Leu Arg Asn Leu Thr Asn Cys Phe Ile Val Ser Leu Ala Ile Thr Asp
    50                  55                  60

Leu Leu Leu Gly Leu Leu Val Leu Pro Phe Ser Ala Ile Tyr Gln Leu
65                  70                  75                  80

Ser Cys Lys Trp Ser Phe Gly Lys Val Phe Cys Asn Ile Tyr Thr Ser
                85                  90                  95

Leu Asp Val Met Leu Cys Thr Ala Ser Ile Leu Asn Leu Phe Met Ile
            100                 105                 110

Ser Leu Asp Arg Tyr Cys Ala Val Met Asp Pro Leu Arg Tyr Pro Val
        115                 120                 125

Leu Val Thr Pro Val Arg Val Ala Ile Ser Leu Val Leu Ile Trp Val
    130                 135                 140

Ile Ser Ile Thr Leu Ser Phe Leu Ser Ile His Leu Gly Trp Asn Ser
145                 150                 155                 160

Arg Asn Glu Thr Ser Lys Gly Asn His Thr Thr Ser Lys Cys Lys Val
                165                 170                 175

Gln Val Asn Glu Val Tyr Gly Leu Val Asp Gly Leu Val Thr Phe Tyr
            180                 185                 190

Leu Pro Leu Leu Ile Met Cys Ile Thr Tyr Tyr Arg Ile Phe Lys Val
        195                 200                 205

Ala Arg Asp Gln Ala Lys Arg Ile Asn His Ile Ser Ser Trp Lys Ala
    210                 215                 220
```

```
Ala Thr Ile Arg Glu His Lys Ala Thr Val Thr Leu Ala Ala Val Met
225                 230                 235                 240

Gly Ala Phe Ile Ile Cys Trp Phe Pro Tyr Phe Thr Ala Phe Val Tyr
                245                 250                 255

Arg Gly Leu Arg Gly Asp Asp Ala Ile Asn Glu Val Leu Glu Ala Ile
            260                 265                 270

Val Leu Trp Leu Gly Tyr Ala Asn Ser Ala Leu Asn Pro Ile Leu Tyr
        275                 280                 285

Ala Ala Leu Asn Arg Asp Phe Arg Thr Gly Tyr Gln Gln Leu Phe Cys
    290                 295                 300

Cys Arg Leu Ala Asn Arg Asn Ser His Lys Thr Ser Leu Arg Ser Asn
305                 310                 315                 320

Ala Ser Gln Leu Ser Arg Thr Gln Ser Arg Glu Pro Arg Gln Gln Glu
                325                 330                 335

Glu Lys Pro Leu Lys Leu Gln Val Trp Ser Gly Thr Glu Val Thr Ala
            340                 345                 350

Pro Gln Gly Ala Thr Asp Arg
        355
```

<210> 4
<211> 15
<212> DNA
<213> Homo sapien

<400> 4
cagctgcrtc tccac 15

<210> 5
<211> 15
<212> DNA
<213> Homo sapien

<400> 5
cttcaggrga ccgtc 15

<210> 6
<211> 15

<212> DNA
<213> Homo sapien

<400> 6
accgtctrag gactg 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapien

<400> 7
actctacygc atgca 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapien

<400> 8
ctgacctrct cctcg 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapien

<400> 9
tcagcctyga ccggt 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapien

<400> 10
atgaagtrta cgggc 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapien

<400> 11
gccatcartg aggtg 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapien

<400> 12
aggatggsgg caatg 15

<210> 13
<211> 15
<212> DNA
<213> Homo sapien

<400> 13
aggaactytt catga 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapien

<400> 14
gggatacagc tgcrt 15

<210> 15
<211> 15
<212> DNA
<213> Homo sapien

<400> 15
ggtcatgtgg agayg 15

<210> 16
<211> 15
<212> DNA
<213> Homo sapien

<400> 16
ctagctcttc aggrg 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapien

<400> 17
tcctcagacg gtcyc 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapien

<400> 18
caggggaccg tctra 15

<210> 19
<211> 15
<212> DNA
<213> Homo sapien

<400> 19
aaactccagt cctya 15

<210> 20
<211> 15
<212> DNA
<213> Homo sapien

<400> 20
gcctggactc tacyg 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapien

<400> 21
tgatcttgca tgcrg 15

<210> 22
<211> 15

<212> DNA
<213> Homo sapien

<400> 22
ctatcactga cctrc 15

<210> 23
<211> 15
<212> DNA
<213> Homo sapien

<400> 23
ggaggccgag gagya 15

<210> 24
<211> 15
<212> DNA
<213> Homo sapien

<400> 24
tcatgatcag cctyg 15

<210> 25
<211> 15
<212> DNA
<213> Homo sapien

<400> 25
cgcagtaccg gtcra 15

<210> 26
<211> 15
<212> DNA
<213> Homo sapien

<400> 26
aggtcaatga agtrt 15

<210> 27
<211> 15
<212> DNA
<213> Homo sapien

<400> 27
ccaccagccc gtaya 15

<210> 28
<211> 15
<212> DNA
<213> Homo sapien

<400> 28
gatgatgcca tcart 15

<210> 29
<211> 15
<212> DNA
<213> Homo sapien

<400> 29
ttctaacacc tcayt 15

<210> 30
<211> 15
<212> DNA
<213> Homo sapien

<400> 30
gtgcacagga tggsg 15

<210> 31
<211> 15
<212> DNA
<213> Homo sapien

<400> 31
ccctcccatt gccsc 15

<210> 32
<211> 15
<212> DNA
<213> Homo sapien

<400> 32
tgttctagga actyt 15

<210> 33
<211> 15
<212> DNA
<213> Homo sapien

<400> 33
aagtgctcat gaara 15

<210> 34
<211> 10
<212> DNA
<213> Homo sapien

<400> 34
atacagctgc 10

<210> 35
<211> 10
<212> DNA
<213> Homo sapien

<400> 35
catgtggaga 10

<210> 36
<211> 10
<212> DNA
<213> Homo sapien

<400> 36
gctcttcagg 10

<210> 37
<211> 10
<212> DNA
<213> Homo sapien

<400> 37
tcagacggtc 10

<210> 38
<211> 10

<212> DNA
<213> Homo sapien

<400> 38
gggaccgtct 10

<210> 39
<211> 10
<212> DNA
<213> Homo sapien

<400> 39
ctccagtcct 10

<210> 40
<211> 10
<212> DNA
<213> Homo sapien

<400> 40
tggactctac 10

<210> 41
<211> 10
<212> DNA
<213> Homo sapien

<400> 41
tcttgcatgc 10

<210> 42
<211> 10
<212> DNA
<213> Homo sapien

<400> 42
tcactgacct 10

<210> 43
<211> 10
<212> DNA
<213> Homo sapien

<400> 43
ggccgaggag 10

<210> 44
<211> 10
<212> DNA
<213> Homo sapien

<400> 44
tgatcagcct 10

<210> 45
<211> 10
<212> DNA
<213> Homo sapien

<400> 45
agtaccggtc 10

<210> 46
<211> 10
<212> DNA
<213> Homo sapien

<400> 46
tcaatgaagt 10

<210> 47
<211> 10
<212> DNA
<213> Homo sapien

<400> 47
ccagcccgta 10

<210> 48
<211> 10
<212> DNA
<213> Homo sapien

<400> 48
gatgccatca 10

<210> 49
<211> 10
<212> DNA
<213> Homo sapien

<400> 49
taacacctca 10

<210> 50
<211> 10
<212> DNA
<213> Homo sapien

<400> 50
cacaggatgg 10

<210> 51
<211> 10
<212> DNA
<213> Homo sapien

<400> 51
tcccattgcc 10

<210> 52
<211> 10
<212> DNA
<213> Homo sapien

<400> 52
tctaggaact 10

<210> 53
<211> 10
<212> DNA
<213> Homo sapien

<400> 53
tgctcatgaa 10

<210> 54
<211> 5335

<212> DNA
<213> Homo sapien

<220>
<221> allele
<222> (3581)
<223> PS1: Polymorphic base G or A

<220>
<221> allele
<222> (3701)
<223> PS2: Polymorphic base G or A

<220>
<221> allele
<222> (3710)
<223> PS3: Polymorphic base G or A

<220>
<221> allele
<222> (4052)
<223> PS4: Polymorphic base C or T

<220>
<221> allele
<222> (4202)
<223> PS5: Polymorphic base G or A

<220>
<221> allele
<222> (4349)
<223> PS6: Polymorphic base C or T

<220>
<221> allele
<222> (4550)
<223> PS7: Polymorphic base G or A

<220>
<221> allele
<222> (4804)
<223> PS8: Polymorphic base A or G

<220>
<221> allele
<222> (5116)
<223> PS9: Polymorphic base G or C

<220>
<221> allele
<222> (5197)
<223> PS10: Polymorphic base C or T

<400> 54
aaggactcca ctgagaagag agtgctccca gggccttgaa aactcaaaac tctgcccaat 60
gggattaaaa aaacaccccc tttctgtcca cctcagaaag cacaagcctc tgcctgaggt 120
gagtacaagg gaggtagaag tgagagagga tggagaacct acctctggaa gacttcctgg 180
aggaggtgac agaggaggcc agggtttgaa aaaagagtaa gagctctcta ggtaaggaaa 240
gaccttccag gcagtgaaaa cagcatgtgc aaaggcacag aatcagaatg ggcctggcac 300
ttttcagaat taaatgccgc agagccagat caacactcca aggcagcagg tgctgtacaa 360
tctaactcgg tacaaagtgc agagagatta ctcctgcttt taggcatcta ggaggactca 420
aggagacagt tacctgagct ggccttgaaa ggagagtggt tgaagtgaat acaagaaaga 480
gtgacacacc aggcataggg gatgcaaagg cagaggaagg agggagaagg tgtgctgcat 540
ctggggacat gctcggttgc ctgagtgtct gtgaaattaa caggcagatg ctggagtgct 600
ggatgagacc ctgttggaaa gacaggttga agtctgatca aagaacgggt attaagtaca 660
ggccaaagga gtttgagctg aatttgaatc caggttctgc cactttctag gggtgtgaca 720
tcacctggcc tctttgagcc tcagtctcct caaccgttaa atggagacat gattttcctc 780
actcgtgggg tcgtgaggat tcagtgatat tttgtttact atgcatggta tcgctttgcc 840
caggaggatc cagcaggaca gttgtgccgt ggttcattaa accttcttcc ccctgcagtt 900
ggcagcttag gttgttttcc attcttgcca ttgcacgtct gccaacggac atccttgtgc 960
acatatcctc atgagagcag gtgcattgcc acgggctggg ctcccgaaag aatggcttcg 1020
atctaagtga cctagagggc catccagtac atattgggaa ttttacaaga agtgaaggag 1080
aggcagttgc tgagtagcat gtatgctgtg tgcgggcagt gcccagtgtg tgcccattgt 1140
gcatgcagag gcatgagaaa cccaggaggc ccccacatac caaacagtca gcagcagtca 1200
cttcctggtg gaggggctca ggtgtgtggt ggctctgatt tttggccctt tctgtacacc 1260
tcagggatat ttggcttact ggaataagcc tgtcatcact ttggaagttt gaaaggacag 1320
tttgagagac tatttttttgt tagctatgat ggtgacaatg acggtgacaa gggtgacatg 1380
attcttcctc acaggcaaag agaacccatt gaaggttgct ccataagggg atgacgtggc 1440
cctctagggc aatccttctg cccacatggg acaggagagg atgggaggga gagagagcgc 1500
cgtacaaaat cccctgccct cgaaaaaggt gaccagaact tacgggagct gggctttaaa 1560
tagtaataca gcagaagtgg tgcttctttg aaaactctac gaagaaagga ctcacagctg 1620
ttttggtcat cattgtcccc ccagcacctg gagtagcatc tgacacattg tgggttttca 1680
aagaaatatg tgttgaaaag gactgaggaa gggaaggagg ggcgggagag gagagaagga 1740
aaggattttc taattataag agaggctggt ccactcagat ttgaagaccc ctcatccaaa 1800
catgaactca aatgtcagcc tgatgccaaa atacctatac ccggctgttg tggtaactca 1860
gaaaaacttt gatgagaaat agggattctc caaaaaaata cccttattgc taccccctcc 1920
aaagttaatt aagtgtggcg gtgagaccca tgctgtttgg gctatttatg taatgaaaag 1980
ctctgctaaa aatagtaccc ggaggttcag ctgccaggac tggtggctga gccagggagc 2040
gtggcagaag gtgatcggga accctgatcc ctgcacgaac acaatcgttg tgattaaatg 2100
atgcttctca tattcaggtt gcaaaccgct ccaggcaaga gcctgggaga ggcttaattt 2160
atgttttatg gagattagct tgcttctgag agggggttcc gccatcaact cgatttctgc 2220
acacactcag agctcaggag attgcctgct ctctgaaaaa ccaggaggag ctggtaggga 2280
aggactgggg cgtggaacag ggcctcagct tgtttctcgc cagttggaat catatacagc 2340
caccactggg agggagcctt ggagattacc tagtgaccag gggctcttat ttaatcaaga 2400
ggcattgcac atgcagttac tttattgacg tagtgtttgc actacttcaa tagtgcaaag 2460
ggaggaaggg aggaaggaga aaccacccta atgcccgtga gagggtgatg tggggatgat 2520

aggataaacc tattcatggt gagcctcccc acagactgtt atgcaaccat tagaaggaat 2580
gagtgagttc tgtccctgct gacctggatg tgggatcacg tgttctgcca agtaagacca 2640
gtaagattta gagatgcgtt cggagcaggg gtggcgtgga tgacaggtga cgaactgccg 2700
attaaatctg gccctctgcc tggctttgta cagcccgtgg ctaagaatgg tttttacatt 2760
tttaaatttt ggaaaaaaat caaaaaaagg agatttttat gacacgtgaa aatgatatga 2820
aatttacatt tcagtgttca ttactgaagt tttgttggag tgcagccacg ctcttctgtc 2880
ggcacgtcat ctgcatagct gcattcgcac tgcaaaggca gagccgagcc gtcacaggct 2940
gtgaggcaga acgtgtttat tacctgaccc ttttctgaaa aagtttgtcg gccaccgcct 3000
tagagtatga tgacattttg ccaaacagca gcaccagata tggttctcag cacttcagag 3060
gagtagattc attacatttg tttattcgtt cattcatttt atcctttatt cgtcagatgc 3120
ttactgtcta tctgtggggc caggcaccag gaaccagcag agaacaaaca gatgcaggcc 3180
ttcttcttgt ggagctgaaa gtctgtgtct agatccttct tttctcagta caagcacttc 3240
cagcctttcc aactctttct cacgagtctc agtgtccagc cccttgcatc ctgcctccct 3300
tctgcctccc cagtatgtat cctgattgcc caagtccctc ccaaaaagag gtatccaaac 3360
tggtcccata cttcctgggg tcttgccaag aggttgagat ctagtagaag gacacatctt 3420
gaattgggtc atgctttcta tctggtttca aggtgcttaa catccaacct ttgccttttc 3480
agctcctgcc ctccactgac tccagagagg gagatcccca gtacttgact ccatcacgca 3540
gatgggagca ggcaccagct atggagaggg atacagctgc rtctccacat gacccatcct 3600
gcatgacacc aaagccaccg ccagacagtg cctcggattc tatgcaaaac ctgggaagcg 3660
gagacctacc ccagccccgg gaggaagcta gctcttcagg rgaccgtctr aggactggag 3720
tttgatccat gaacctggct tcgaggcctt gcttttctct cttcttcatt catattcatt 3780
cccaacacct tagaaggtgt tgcttaattt atttctagaa aagcagccca gagtcagtca 3840
ttgaagcctt ccccaccccc tggccaaaaa aaaaaaaaaa aaaaaactgg acacattttg 3900
gatctgttgg gagcttggag tccagtggtt ggcatagttg tcacattggg agcagagaag 3960
aagcaaccag gggccctgat caggggactg agccgtagag tcccaggatg gcacccaatg 4020
gcacagcctc ttccttttgc ctggactcta cygcatgcaa gatcaccatc accgtggtcc 4080
ttgcggtcct catcctcatc accgttgctg gcaatgtggt cgtctgtctg gccgtgggct 4140
tgaaccgccg gctccgcaac ctgaccaatt gtttcatcgt gtccttggct atcactgacc 4200
trctcctcgg cctcctggtg ctgcccttct ctgccatcta ccagctgtcc tgcaagtgga 4260
gctttggcaa ggtcttctgc aatatctaca ccagcctgga tgtgatgctc tgcacagcct 4320
ccattcttaa cctcttcatg atcagcctyg accggtactg cgctgtcatg gacccactgc 4380
ggtaccctgt gctggtcacc ccagttcggg tcgccatctc tctggtctta atttgggtca 4440
tctccattac cctgtccttt ctgtctatcc acctggggtg gaacagcagg aacgagacca 4500
gcaagggcaa tcataccacc tctaagtgca aagtccaggt caatgaagtr tacgggctgg 4560
tggatgggct ggtcaccttc tacctcccgc tactgatcat gtgcatcacc tactaccgca 4620
tcttcaaggt cgcccgggat caggccaaga ggatcaatca cattagctcc tggaaggcag 4680
ccaccatcag ggagcacaaa gccacagtga cactggccgc cgtcatgggg gccttcatca 4740
tctgctggtt tccctacttc accgcgtttg tgtaccgtgg gctgagaggg gatgatgcca 4800
tcartgaggt gttagaagcc atcgttctgt ggctgggcta tgccaactca gccctgaacc 4860
ccatcctgta tgctgcgctg aacagagact tccgcaccgg gtaccaacag ctcttctgct 4920
gcaggctggc caaccgcaac tcccacaaaa cttctctgag gtccaacgcc tctcagctgt 4980
ccaggaccca aagccgagaa cccaggcaac aggaagagaa acccctgaag ctccaggtgt 5040
ggagtgggac agaagtcacg gccccccagg gagccacaga caggtaatag ccctagccat 5100
tggtgcacag gatggsggca atgggagggg atgctactga tgggaatgat taagggagct 5160
gctgtttagg tggtgctggt ttatgttcta ggaactyttc atgagcactt tgtaaacacc 5220
ctcttgctta atcctcccaa cggcccccaa aggtagaact tagctccctt ttaaaaggag 5280
cacattaaaa ttctcagagg acttggcaag ggccgcacag ctggggcatg aattc 5335

Am J Physiol Endocrinol Metab 286: E252–E260, 2004.
First published October 21, 2003; 10.1152/ajpendo.00424.2003.

Identification of a novel mitochondrial protein ("mitoNEET") cross-linked specifically by a thiazolidinedione photoprobe

Jerry R. Colca, William G. McDonald, Daniel J. Waldon, Joseph W. Leone, June M. Lull, Carol A. Bannow, Eric T. Lund, and W. Rodney Mathews
Pharmacia Corporation, Kalamazoo, Michigan 49001

Submitted 21 September 2003; accepted in final form 9 October 2003

Colca, Jerry R., William G. McDonald, Daniel J. Waldon, Joseph W. Leone, June M. Lull, Carol A. Bannow, Eric T. Lund, and W. Rodney Mathews. Identification of a novel mitochondrial protein ("mitoNEET") cross-linked specifically by a thiazolidinedione photoprobe. *Am J Physiol Endocrinol Metab* 286: E252–E260, 2004. First published October 21, 2003; 10.1152/ajpendo.00424.2003.— Thiazolidinediones address underlying causes of type 2 diabetes, although their mechanism of action is not clearly understood. The compounds are thought to function as direct activators of the nuclear receptor PPARγ (peroxisome proliferator-activated receptor-γ), although pioglitazone, the weaker agonist of the two thiazolidinediones now in clinical use, seems to have more useful effects on circulating lipids. We have used tritiated pioglitazone and a photoaffinity cross-linker to identify a novel binding site in mitochondria. A saturable binding site for [^{3}H]pioglitazone was solubilized from the membranes with CHAPS and migrated as a large complex by size exclusion chromatography. The binding correlated with a <17-kDa protein (m17), marked by a photoaffinity cross-linker, in both subcellular location and selectivity of competition by analogs. The protein was isolated and identified by mass spectrometry analysis and NH_2-terminal sequencing. Three synthetic peptides with potential antigenic properties were synthesized from the predicted nontransmembrane sequence to generate antibodies in rabbits. Western blots show that this protein, which we have termed "mitoNEET," is located in the mitochondrial fraction of rodent brain, liver, and skeletal muscle, showing the identical subcellular location and migration on SDS-PAGE as the protein cross-linked specifically by the thiazolidinedione photoprobe. The protein exists in low levels in preadipocytes, and expression increases exponentially in differentiated adipocytes. The synthetic protein bound to solid phase associated with a complex of solubilized mitochondrial proteins, including the trifunctional β-oxidation protein. It is possible that thiazolidinedione modification of the function of the mitochondrial target may contribute to lipid lowering and/or antidiabetic actions.

insulin sensitizers; thiazolidinediones; photoaffinity probes; novel site of action; mitoNEET

THIAZOLIDINEDIONES WERE FIRST DESCRIBED by the Takeda Company as lipid-lowering, antidiabetic agents in the early 1980s (6, 14, 22). Treatment of animals resulted in improvement of insulin action in all of the target tissues, and various lines of evidence suggested that the pharmacology of these compounds was secondary to improved insulin sensitivity (8, 21). The work of Kletzien and colleagues (18, 26, 27) provided the biochemical basis for the hypothesis that these agents produce their effect in vitro by direct interaction with the nuclear receptor peroxisome proliferator-activated receptor-γ (PPARγ) (31, 43). The correlation of these studies in vitro on the differentiation of adipocytes and antidiabetic action extended the hypothesis to suggest that the lipid-lowering and antidiabetic actions of these molecules were secondary to direct activation of PPARγ (40, 43, 48).

Although there have been considerable attempts to exploit this hypothesis (36), and structural information regarding the putative target is available (46), there are no superior drugs on the horizon. Furthermore, there are some inconsistencies in the hypothesis that direct activation of PPARγ might explain all of the pharmacology. Thus not all PPARγ activators have antidiabetic actions (30), and some potent antidiabetic analogs are not good PPARγ activators (5, 37, 42). Partial reduction in PPARγ expression results in increased insulin sensitivity (33), and naturally occurring mutations also do not offer a clear picture (e.g., 17, 19, 44). Finally, in humans, pioglitazone (Actos), a weaker activator of PPARγ (31), generally produces a greater reduction in circulating lipids (reduction in triglycerides and total cholesterol/HDL cholesterol) than does the more potent activator rosiglitazone (Avandia) (2, 15, 24, 29).

Given the lack of certainty about the mechanism of action of this important class of drugs (28, 35), we have explored the possibility of an alternate site of binding of the thiazolidinediones. We have approached this problem by using high specific activity tritiated pioglitazone and a structurally related iodinated photoaffinity probe.

MATERIALS AND METHODS

Probes. [^{3}H]pioglitazone was generated by tritium exchange from the dibromo derivative at 31.7 Ci/mmol and was stored in methanol. PNU-101074 was synthesized by coupling a carboxylic acid analog of pioglitazone (PNU-91323) to a *p*-azido-benzyl group containing ethylamine (Fig. 1). The purified compound was iodinated, carrier-free, with Iodogen (Pierce), and ^{125}I-PNU-101074 was purified on a C_{18} column and stored in the dark. These radioactive probes and competitor compounds are shown in Fig. 1. Competitor compounds, to show the selectivity of binding/cross-linking, were made as previously described (45).

Membrane isolation. Crude rat liver, skeletal muscle, and brain mitochondrial-enriched fractions were prepared as follows. Briefly, Sprague-Dawley rats were anesthetized, and hindleg muscle, liver, and whole brain were removed to a cold buffer, MLB (in mM: 225 sucrose, 6 K_2HPO_4, 5 $MgCl_2$, 20 KCl, and 2 EGTA, pH 7.4). Tissues were chopped, rinsed, and homogenized with a polytron (setting 7; 3×, 15 s) in 5 volumes of MLB. After removal of the unbroken cells and nuclei (750 *g*), the mitochondrial-enriched fraction was collected at 15,800 *g* for 5 min. The loose pellet was discarded, and the dense central pellet was resuspended in MLB and recollected at 11,800 *g* for 10 min. The final pellets were resuspended in 50 mM Tris (pH 8) at 5–8 mg/ml total protein and frozen at −80°C.

Address for reprint requests and other correspondence: J. R. Colca, Pfizer St. Louis, 700 Chesterfield Pkwy, Chesterfield, MO 63198 (E-mail: jerry.r.colca@pfizer.com).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

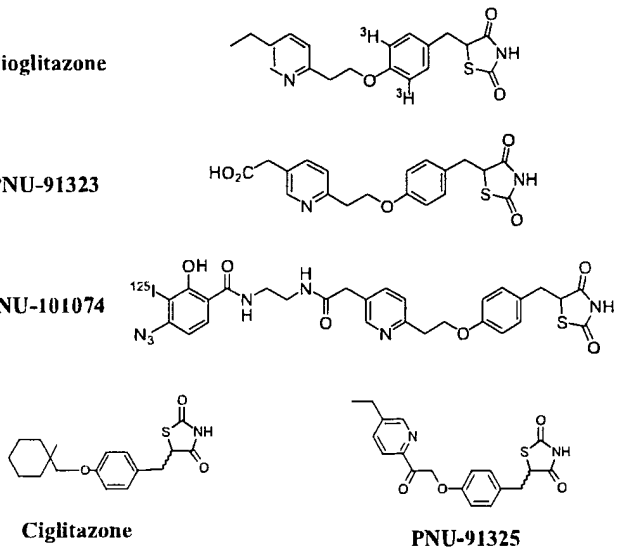


Fig. 1. Probes used to identify binding site. Pioglitazone was tritiated to ~31 Ci/mmol, as described in text. The photoaffinity probe, PNU-101074, was prepared from PNU-91323 and iodinated (~2,000 Ci/mmol) as described. Ciglitazone and PNU-91325 were used as competitor compounds.

Bovine brain mitochondria were harvested from steer brains. The rinsed brains were homogenized in fractionation buffer (250 mM sucrose, 50 mM Tris, pH 8.0, containing 1 mg/ml pepstatin A, 5 mg/ml leupeptin, 10 mg/ml bacitracin, and 0.1 mM PMSF). After removal of nuclei at 2,250 rpm in a Sorval SS-34 rotor, the mitochondrial pellet was harvested at 20,000 *g* (13,000 rpm in a Sorval SS-34 rotor) and further enriched by sucrose density centrifugation. Membrane fractions were collected from the top of the 1.18 and 1.20 density bands, resuspended in 50 mM Tris, and collected by centrifugation. The fractions ("B3/B4") were stored at −80°C until use. Lower-density bands were collected as controls. Succinate cytochrome *c* reductase was used as the mitochondrial marker (12), and protein was measured by the bicinchoninic acid method (Pierce).

3T3-L1 preadipocytes were grown and differentiated as previously described (41). Preadipocytes or fully differentiated adipocytes (11 days) were lifted off the plate with trypsin, and the cells were washed with Dulbecco's PBS and homogenized with 15 strokes with a motor-driven Teflon homogenizer. For the differentiated adipocytes, both floating and pelleted cells were included. After removal of the low-speed (750 *g*) pellet, the second pellet was collected at 18,000 *g* for 10 min. The pellets were resuspended in 50 mM Tris (pH 8), equalized for protein content, and used for cross-linking and Western blots as we will describe.

Binding and cross-linking assays. [^{3}H]pioglitazone binding was conducted using intact mitochondrial fractions and also fractions solubilized in 20 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS). Intact membrane solutions were thawed and diluted to a concentration of 1 mg/ml in 50 mM Tris, pH 8.0. The reactions were conducted in 1.5-ml polypropylene microfuge tubes. The assay volume was 400 μl, containing 100 μl of membranes, 100 μl of 0.4% gamma globulin, 100 μl of 4% DMSO with or without the indicated concentration of competitor, and 100 μl of [^{3}H]pioglitazone (0.2 μCi). The tubes were mixed and incubated at ambient temperature on a rocker for 60 min. Bound counts were detected after sedimentation of membranes (18,000 *g*, 5 min). For study of the solubilized binding, membrane stocks were solubilized in 20 mM CHAPS, and bound counts were separated from free ones with dextran-coated charcoal (5% charcoal/0.5% dextran in 50 mM Tris, pH 8.0). All binding assays were conducted in triplicate.

Cross-linking reactions were carried out in a final volume of 200 μl, containing 100 μl of membranes (adjusted for total protein), 50 ml of 4% DMSO with or without competing thiazolidinedione (TZD; e.g., 100 μM PNU-91325, 25 μM final concentration), and 50 μl of carrier-free ^{125}I-PNU-101074 (0.1–0.2 μCi/tube). Specific cross-linking was defined as labeling that was prevented in the presence of competing active TZD (29). An appropriate amount of ^{125}I-PNU-101074 in acetonitrile was dried under vacuum in the dark immediately before use. The reactions were incubated for 15 min at room temperature and stopped by exposure to UV light in open tubes (180,000 μJ in a Stratalinker). The cross-linked samples were then rinsed with 50 mM Tris (pH 8.0) after centrifugation at 18,000 *g* for 5 min. The rinsed pellets were resuspended in 100 μl of 50 mM Tris.

Purification and identification of cross-linked protein. Optimal selective solubilization of the specifically cross-linked protein (m17) was obtained by bringing the resuspended pellet to 1% Triton X-114. After rocking at room temperature for 5 min, the bulk of the cross-linked m17 was recovered in the supernatant after centrifugation at 18,000 *g* for 15 min. The cross-linked m17 also remained in the supernatant after centrifugation at 450,000 *g* in a Beckman TLA 100 ultracentrifuge for 30 min. This removed many of the contaminating proteins.

Triton X-114 was removed from the sample by precipitation and flotation of the detergent with ammonium sulfate. The addition of equal volumes of 1.5 M ammonium sulfate to the Triton X-114 solution precipitated the proteins, leaving the detergent in solution. The precipitation procedure was repeated 3 times to maximize the yield of precipitated protein. Special care was taken to ensure that the protein precipitate did not float on the Triton-containing supernatant. The precipitated protein was concentrated for direct separation on SDS-PAGE gels (10–20% or 18% Tris-glycine) or for HPLC.

HPLC was conducted by using both UV scanning diode array detector and an in-line gamma-C flow cell (Packard). Purification of the m17 protein was achieved with a Phenomenex Synergi MAX-RP C_{12} column (4.6 × 250 mm, 4 μm). The guard column was a 4 × 2.0-mm RP-1 SecurityGuard Cartridge (Phenomenex). The selections of the column and guard columns were made after considerable examination of standard protein columns, which gave no appreciable yield of the target protein. Samples were eluted with a programmed gradient elution starting with 70% *solution A* [water-0.05% trifluoracetic acid (TFA), vol/vol] and 30% *solution B* (Can-0.05% TFA, vol/vol). The gradient was held at 30% *B* for the first 15 min; *B* was then increased from 30 to 55% over 30 min and then increased to 80% over 15 min. Flow rates were fixed at 1 ml/min.

The specifically cross-linked proteins were excised from electrophoretic gels and were reduced, alkylated, and digested in situ with modified porcine trypsin (Promega) by use of a DigestPro robot (ABIMED). Briefly, protein gel spots were placed in reaction vials and secured in a Peltier heating/reaction block. Digested peptides were extracted with 60% acetonitrile-5% formic acid. Peptide extracts were placed in a Speed-Vac centrifuge until dry and were reconstituted in 10 μl of 5% formic acid in water.

NanoLC tandem mass spectrometry analysis (nanoLC-MS/MS) was performed on a Micromass Qtof ultima instrument coupled to a Micromass CapLC. Typically, 5 μl from a total sample amount of 5.5 μl were injected and preconcentrated using column switching. An auxiliary pump was used to preconcentrate and desalt samples on a C_{18} Pepmap precolumn (0.3 × 5 mm) by delivering 0.1% formic acid at 20 μl/min. After desalting, the precolumn was switched in-line with the analytical column (75-mm-ID C_{18} Pepmap, LC Packings) and eluted at 300 nl/min with a gradient of 0.1% formic acid in water and 0.1% formic acid containing 90% acetonitrile directly into the Qtof. Tandem MS data were acquired and processed by Micromass MassLyxn software. Nanospray MS/MS data were used to identify proteins by comparing the experimental data with predicted data derived from protein and DNA databases. Tandem MS data were searched against

the NCBInr protein database with MASCOT (Matrix Science) programs maintained on the SAM Chemistry MS lab NT server.

A procedure was developed to elute the m17 after SDS-PAGE with a series of rehydration and drying steps. The individual lane bands were cut out, and groups of 10 gel slices were placed in microfuge tubes, rehydrated in 0.5 ml of water, and nutated for 1 h at room temperature. The eluted material was retained, and the gel slices were then dried in a Speed-Vac centrifuge; this procedure was repeated two additional times. The pooled, eluted m17 was concentrated and resolved on SDS-PAGE before MS/MS identification or used for generation of CNBr fragments in 70% formic acid (overnight at room temperature). CNBr fragments were released by water elution. No further recovery occurred by electroelution of these gels. The samples were finally concentrated, run on 18% Tris-glycine gels (Invitrogen), and blotted to Immobilon-Psq (Millipore). The blots were stained with 0.1% Coomassie R-250, destained, air dried, and exposed to Biomax MS film at −80°C. Amino-terminal sequencing was performed by automated Edman degradation on an Applied Biosystems model 492 Procise cLC protein sequencer.

Generation of antibodies and Western blotting. The protein identified from cross-linking with PNU-101074 was evaluated for potential antigenic peptides with a set of computer programs developed in-house by F. J. Kezdy and R. A. Poorman. Three peptides were chosen and synthesized on an Applied Biosystems 433A peptide synthesizer by use of a HBTU/NMP protocol with the Fmoc (9-fluorenylmethoxycarbonyl) group as the NH_2- or amino-protecting group. The crude peptides were precipitated, purified by reverse-phase HPLC (Vydac C_{18}, 22 × 250 mm, 10 μm), and characterized by open access electrospray mass spectrometry.

Peptides A (CGGKAMVNLQIQKDDPKVV-OH), *B* (KVVHAFDMGDLGDKAVWC-OH), and *C* (CGGNEETGDNVGPLIIKKKET) (compare with sequence in Fig. 5) were sent to Covance Research Products (Denver, PA) for conjugation to keyhole limpet hemocyanin and immunization of rabbits. Serum was tested against a spotted concentration-dose curve (0.01–10 mg) against all peptides. Positive reactions were obtained from the first bleed onward for *peptides A* and *B*. *Peptide C* did not elicit an immune response. Antisera to *A* or *B* did not cross-react to any of the other peptides. Western analysis was conducted by running protein samples on reducing 18% Tris-glycine gels and blotting to polyvinyldifluoride membranes. The blotted membranes were incubated with a 1:30,000 dilution of anti-mitoNEET *peptide B*. Detection of the immunoreactive bands was determined by incubating with a 1:50,000 dilution of alkaline phosphatase-conjugated monoclonal anti-rabbit IgG (Sigma no. A-2556) followed by development with BCIP/NBT Blue Liquid Substrate (Sigma no. B-3804). Anti-prohibitin (1:400 dilution; Research Diagnostics) was used as a mitochondrial protein marker (23). The developed blots were dried and exposed to Biomax MS film.

Association of mitochondrial proteins with the synthetic target. The 108-amino acid protein identified as the putative target of TZD cross-linking was synthesized by solid phase to contain an NH_2-terminal biotin for attachment of streptavidin agarose (4%) beads (Sigma, S-1638). Peptide with or without an NH_2-terminal biotin extension (25 μg/50 μl bead suspension) was incubated in a total volume of 250 μl for 1 h at 4°C, followed by addition of 50 μl of 1 mM biotin and 200 μl of the solubilized mitochondrial preparation and incubation for an additional 2 h. The beads were then washed 6 times with 1 ml of 10 mM Tris, 150 mM NaCl, and 0.4 mM CHAPS. Proteins were eluted with a minimal volume (60 μl) of 0.1 M glycine and 150 mM NaCl, pH 2.3. There was no binding of solubilized mitochondrial proteins to beads alone or beads plus the peptide not containing the NH_2-terminal biotin.

RESULTS

A saturable binding site for [^{3}H]pioglitazone was found in crude mitochondrial membranes from bovine brain (Fig. 2*A*). Similar results were obtained with crude mitochondrial membranes from rat liver, skeletal muscle, and differentiated 3T3-L1 adipocytes. In these studies, half-maximal binding occurred between 0.1- and 1-μm pioglitazone. No specific binding was measured in any other cellular fractions (see below). The binding activity was solubilized by CHAPS. The soluble binding had similar characteristics to the binding in the intact membranes and showed the same specificity with competition by unlabeled pioglitazone, but less competition by ciglitazone (Fig. 2*B*), which itself is considered to have little activity (20).

Attempts to isolate the soluble [^{3}H]pioglitazone-binding fraction met with little success. The activity migrated as a large complex in the void volume of size-exclusion columns and as a single peak on ion exchange, but the activity was rapidly lost, especially when the samples were not kept at a pH close to 8 (not shown). A photoaffinity probe was made to mark the site of attachment of the TZD for further characterization.

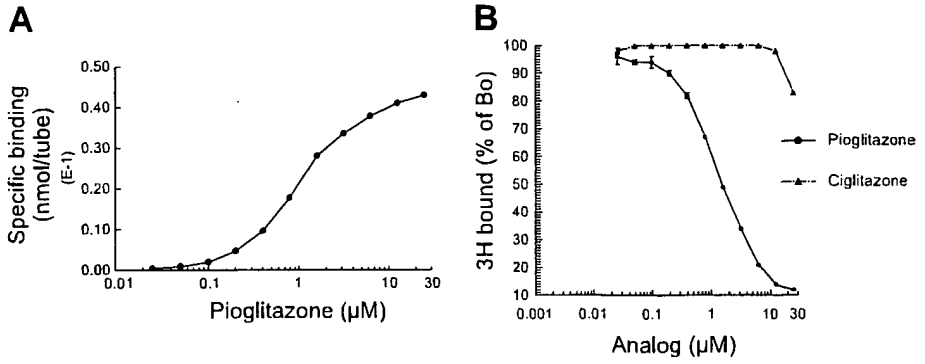


Fig. 2. Binding of [^{3}H]pioglitazone to mitochondrial membranes. *A*: intact mitochondria from bovine brain were incubated with the concentration of pioglitazone shown on abscissa. Specific binding (less the nonspecific binding, counts at 100 μM excess cold pioglitazone) was measured as described in text and is shown on the ordinate (mean of triplicates). Similar results were obtained from other tissues on ≥3 occasions. *B*: binding of [^{3}H]pioglitazone was measured in solubilized bovine brain mitochondria, as described in text. Decrease in specific binding was measured at increasing concentrations of unlabeled pioglitazone and ciglitazone (abscissa). Data on ordinate are percentages of specific binding (B_o) in the absence of added competitors. Values are means ± SE of triplicates. Similar results were obtained from mitochondria from other tissues. Each experiment was repeated on ≥3 occasions.

As shown in Fig. 3*A*, a single specifically cross-linked band (competed by the active TZD, PNU-91325) marked by the arrow (m17), was found in the crude mitochondrial fraction obtained from a variety of tissue sources. Competition for cross-linking of this band by both pioglitazone and ciglitazone was similar to the competition for [^{3}H]pioglitazone binding (Fig. 3*B*). Furthermore, both [^{3}H]pioglitazone binding (Fig. 3*C*) and cross-linking of the m17 protein (Fig. 3*D*) correlated with the mitochondrial marker succinate cytochrome *c* reductase in subcellular fractions. Similar results were obtained for rat skeletal muscle and liver (not shown).

The cross-linked band could not be resolved by two-dimensional gel electrophoresis (not shown), a technique that has proven successful for other proteins identified by this technique (9). The specifically cross-linked band was purified from bovine brain and rat liver mitochondria by two approaches. Cross-linked mitochondria were solubilized with 1% Triton X-114, also resulting in a partial enrichment with respect to total protein. Further enrichment and concentration of m17 were accomplished by precipitating the solubilized cross-linked protein with 0.75 M ammonium sulfate (AS). This was the optimal concentration of AS that allowed precipitation of the protein while keeping the Triton in solution. Concentration and removal of the Triton X-114 were essential for optimal separation by HPLC (see below).

The concentrated m17 was separated by HPLC, as described in MATERIALS AND METHODS. Identical results were obtained from either fresh rat liver mitochondrial samples or bovine brain mitochondrial fractions, suggesting that a similar target protein was involved. A representative pattern of the separation by HPLC is shown in Fig. 4*A*. Identification of the radioactive peak was simplified by the in-line radiometric detector (Fig. 4*A*, *bottom right*). The m17 peak eluted at ~30 min under these conditions at ~55% acetonitrile. Parallel runs with samples from cross-linking incubations that contained the competitor PNU-91325 lacked this peak (not shown). SDS-PAGE together with autoradiography demonstrated that this method provides a successful purification of the specifically PNU-101074-cross-linked protein (Fig. 4).

The m17 cross-linked protein was also concentrated in high yield by a water elution procedure from unfixed, unstained gels (Fig. 4*B*). For this approach, 80 individual tubes were cross-linked with or without PNU-91325, solubilized with Triton X-114, concentrated by AS precipitation, and then subjected to SDS-PAGE on 18% Tris-glycine gels that were not fixed or stained. The bands of interest were marked and cut out as described in MATERIALS AND METHODS. Figure 4*B* shows an autoradiogram of a representative gel before and after the band of interest was cut out for water elution of the cross-linked m17 band. Reexposure of these gels confirmed that the center of the

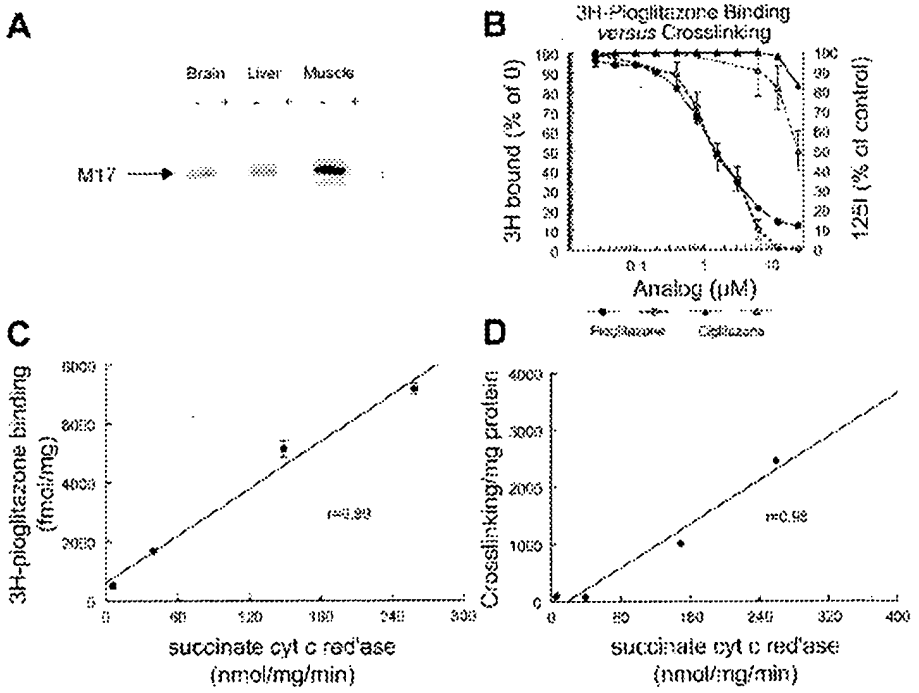


Fig. 3. Cross-linking with ^{125}I-PNU-101074. *A*: a representative autoradiogram of specific cross-linking of a protein band beneath the 17-kDa marker. Intact mitochondrial fractions from rat brain, liver, and skeletal muscle were cross-linked without (−) or with excess unlabeled thiazolidinedione competitor (+). Samples were separated on 18% polyacrylamide gels that were dried and exposed to film. Arrow, specifically cross-linked band. *B*: comparison of the relative ability of pioglitazone (circles) and ciglitazone (triangles) to compete for [^{3}H]pioglitazone binding (solid symbols/solid line) and ^{125}I-PNU-101074 cross-linking (open symbols/dotted line) of the <17-kDa protein. Values are means ± SE for triplicate measurements. These data are from an intact bovine brain mitochondrial preparation and are representative. Similar data were generated using crude rat skeletal muscle mitochondria (not shown). *C*: correlation of specific binding of [^{3}H]pioglitazone binding (fmol/mg, ordinate) vs. specific activity of mitochondrial marker succinate cytochrome *c* reductase (cyt c red'ase, abscissa). These data are from sucrose density gradient bands derived from a 20,000-*g* pellet of bovine brain, as described in text. *D*: the same subcellular fractions described in *C* were used for cross-linking with ^{125}I-PNU-101074. Counts/mg protein (ordinate) were correlated with the mitochondrial marker succinate cytochrome *c* reductase (abscissa). Correlations shown were repeated in independent experiments.

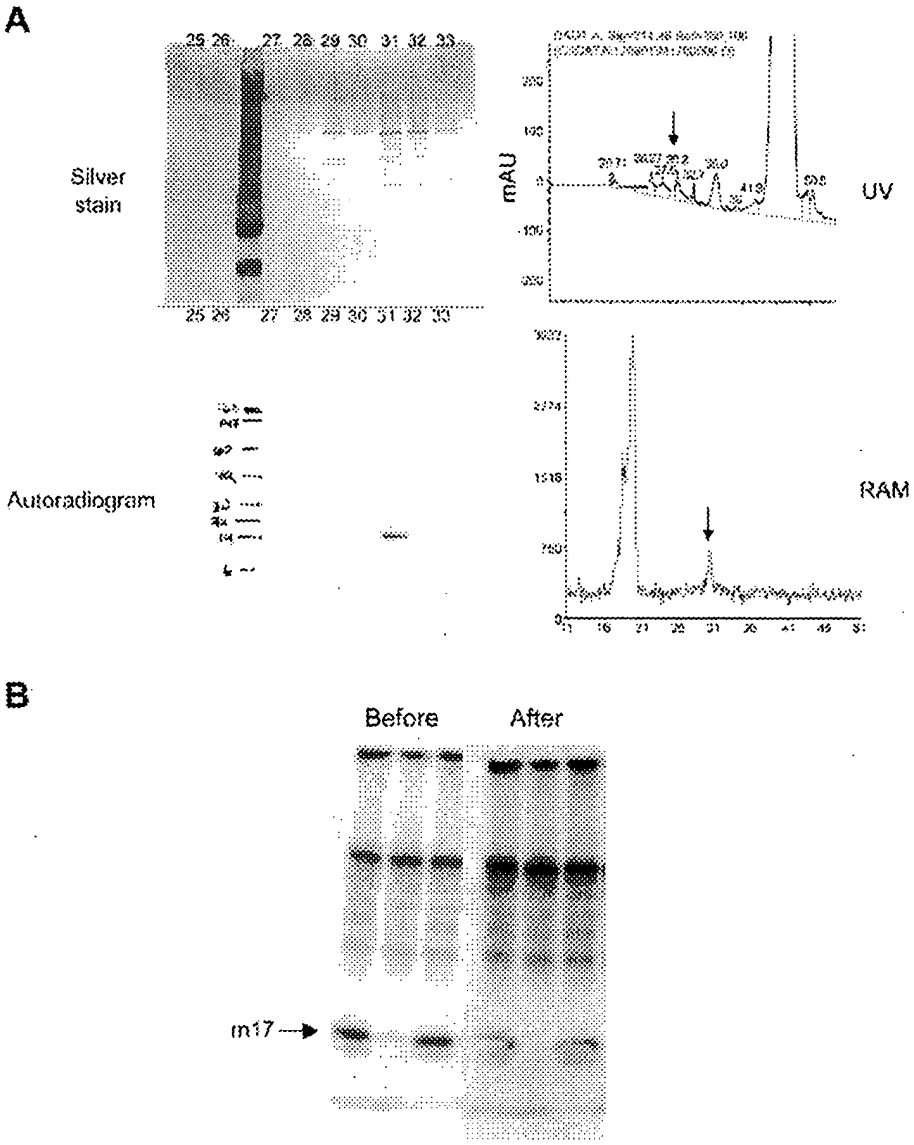


Fig. 4. Purification of the ^{125}I-PNU-101074-cross-linked protein. Specifically cross-linked m17-kDa protein was selectively solubilized from bovine brain or rat liver mitochondria. After removal of detergent, samples were subjected to HPLC and/or electrophoresis, as described in text. *A*: HPLC profiles of UV and ^{125}I (RAM) are shown at *right*. *Left*: a representative silver-stained gel (*top left*) and respective autoradiogram (*bottom left*) of the fractions containing the target protein. *B*: representative autoradiograms of unfixed, unstained gels showing successful excision of the center of the gel containing the specifically cross-linked protein. Three lanes are shown (center lane was cross-linked in the presence of unlabeled competitor) before (*left*) and after (*right*) removal of the center of the band with a scalpel.

specific band had been excised. This procedure produced the highest yield of m17 cross-linked protein.

Purified m17 bands from unfixed, unstained 18% Tris-glycine gels were rinsed and processed for proteomic identification, as described in MATERIALS AND METHODS. Preparations from both rat liver mitochondria and bovine mitochondrial fractions identified the same tryptic peptides shown in bold in Fig. 5*A*. These sequences are annotated "similar to hematopoietic stem/progenitor cells protein MDS029" ID gi 15488774: (BC013522). The mouse protein sequence is shown. The predicted human and mouse proteins are virtually identical.

We attempted to confirm the identification by NH_2-terminal sequencing. Sequencing of the intact protein was unsuccessful, suggesting that the NH_2 terminus might be blocked. Gel digestion with CNBr generated a <6-kDa cross-linked fragment. The blot showing the intact and CNBr-digested fragment is shown in Fig. 5*B*. Partial sequence data were obtained from this fragment supporting the MS/MS identification of the labeled protein (Fig. 5*B*).

We next sought to generate antibodies against this protein. Three peptides from the predicted non-membrane-spanning region were selected and synthesized. Antisera generated against both *peptide A* and *peptide B* recognized the m17 on Western blots; however, the greatest reactivity was with the serum generated from rabbits immunized with *peptide B*. Figure 6 demonstrates a representative Western blot of cross-linking reactions with sucrose density fractions obtained from bovine brain. As predicted from previous results, the cross-linked band was enriched in the higher-density protein bands that are enriched in mitochondrial marker (in agreement with results shown in Fig. 3, *C* and *D*). After SDS-PAGE, representative gels for these samples were transferred to membranes for Western blots using preimmune (Fig. 6, *top left*), anti-*peptide B* (Fig. 6, *top center*), or anti-prohibitin, a known mitochondrial protein (23) (Fig. 6, *top right*). Prohibitin and m17 immunoreactivity were in the same fractions, and the m17 staining was overlaid by the TZD-specific cross-linking (Fig. 6, *bottom*) representative panels.

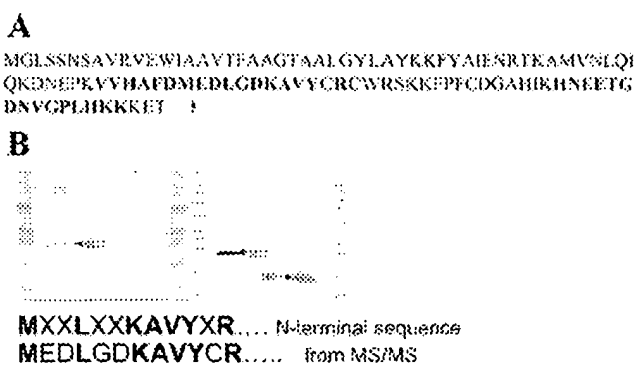


Fig. 5. Identification of the ^{125}I-PNU-101074-cross-linked protein. *A*: sequence identified by nanospray LC mass spectroscopy (MS/MS) from both bovine brain and rat liver mitochondria. The 2 tryptic fragments sequenced by MS/MS are shown in bold within the entire sequence of the target protein. This result was repeated on 3-4 occasions for each tissue. *B*: a Coomassie blue-stained blot of concentrated protein is shown together with autoradiogram image of blot. Gels contained the intact protein (m17) and the <6-kDa CNBr-digested fragment that still contained the cross-linked probe (m6). NH_2-terminal sequence data are shown (X = cycle residue not identified) and compared with sequence predicted by data described in *A*. Matching amino acids are in bold, and arrow in *A* shows point of CNBr cleavage.

The antibody also recognized a protein band of the same size as the specifically cross-linked band in rat liver and skeletal muscle mitochondrial fractions; other cell fractions were negative (not shown).

A further correlation of the m17 protein by Western blot analysis and cross-linking by the TZD probe is shown in Fig. 7. As can be seen in the autoradiograms depicting the TZD-specific cross-linking, there is little cross-linking (Fig. 7, *B* and *D*, *lanes 1–4*) and no measurable immunostaining in the Western blots (Fig. 7*C*) in the crude membranes from preadipocytes (*lanes 1–4*), whereas both end points are found in membranes from the differentiated cells (*lanes 5–8*). Overexposure of the autoradiograms did demonstrate very low levels of cross-linking in the preadipocytes (not shown).

Although antibodies against the m17 peptides recognized the intact protein, they were not useful for immunoprecipitation purposes (not shown). Because the tritiated pioglitazone-binding activity in solubilized membranes migrated as a large complex (see RESULTS), we assumed that the m17 might associate with other mitochondrial proteins and that this might provide some hint as to its function. The 108-amino acid protein was synthesized by a solid-phase approach with an NH_2-terminal biotin and fastened to Streptavidin beads, as described in MATERIALS AND METHODS. Solubilized mitochondrial preparations were then incubated with the beads with or without the attached synthetic protein. Figure 8 shows an example of a silver-stained gel containing the contents released from these beads. In the absence of peptide (−, beads only), no bound proteins were eluted when the pH of the washing solution was reduced. In the presence of the peptide (+), a

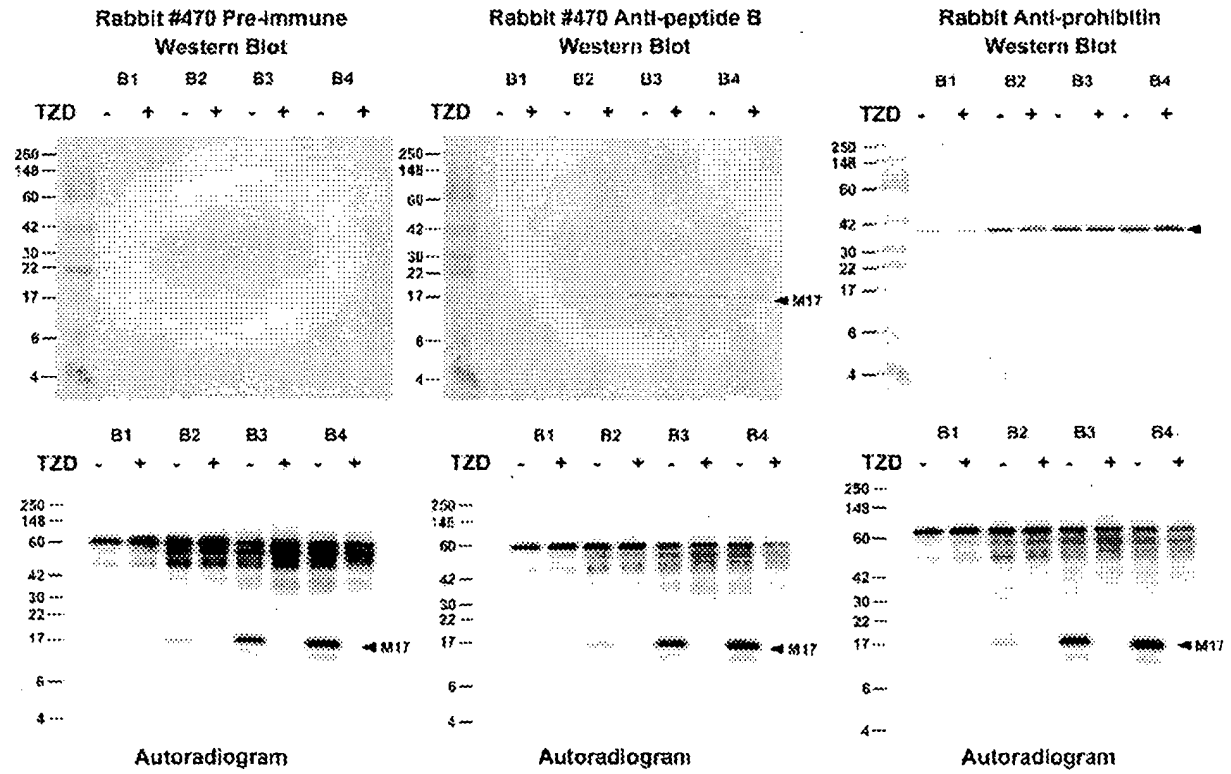


Fig. 6. Western blot for putative target of ^{125}I-PNU-101074 cross-linking. Cross-linking reactions were conducted on sucrose density gradient bands from subcellular fractions of bovine brain mitochondria, as described in text and in Fig. 3, *C* and *D*. Samples were subjected to electrophoresis on 3 polyacrylamide gels, blotted, and then incubated with preimmune serum (*left*), antibodies to m17 *peptide B* (*center*), or antibodies to mitochondrial protein prohibitin (*right*). Western blots (*top*) were conducted and developed as described in text. Blots were then exposed to film, and respective autoradiogram images are shown at *left*. The same results were obtained from a 2nd experiment.

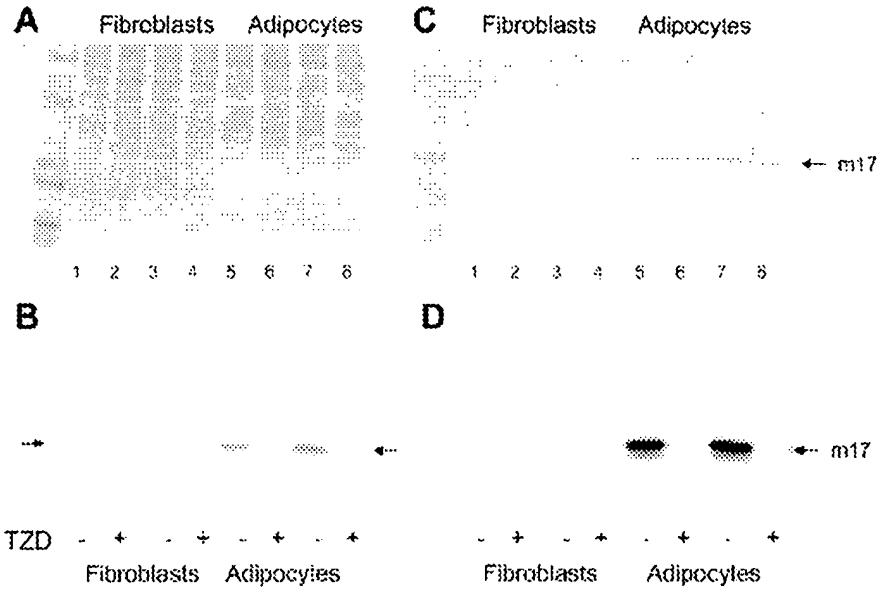


Fig. 7. Western blot for putative target in 3T3-L1 fibroblasts and adipocytes. Pellets (18,000 *g*) were made from 3T3-L1 preadipocytes (fibroblasts) and fully differentiated adipoctyes, as described in text, and equalized for total protein. Cross-linking reactions were run without (−) or with (+) excess unlabeled TZD. Samples without (odd numbers) and with (even numbers) competitors were conducted in duplicate for preadipoctye (fibroblast) (*lanes 1–4*) and adipocyte (*lane 5–8*) pellets. Samples were then subjected to electrophoresis and stained for protein with Coomassie blue (*A*) or used for Western blots with antisera to *peptide B* (*C*). Respective autoradiograms are shown beneath stained gel (*B*) and blot (*D*).

subset of proteins was eluted from beads incubated with solubilized mitochondria from brain, skeletal muscle, or liver. The protein bands were cut from the gel and identified by nanoLC-MS/MS, as described in MATERIALS AND METHODS. A representative data set obtained from a solubilized rat liver mitochondria preparation is shown in Table 1. The subset of proteins collected by m17 from mitochondria from various sources included complex 3, ATP synthase, and enzymes involved in β-oxidation of fatty acids.

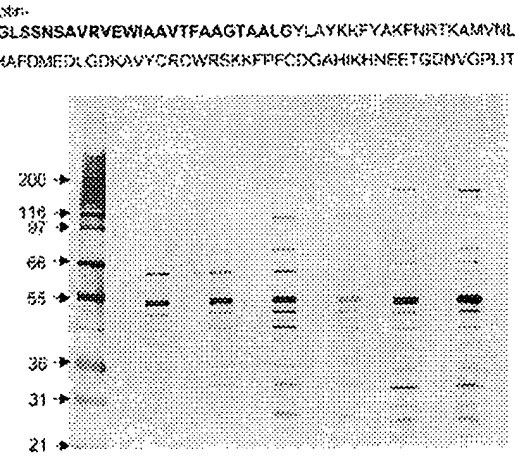


Fig. 8. Binding of solubilized mitochondrial proteins to synthetic m17 kDa target protein. Synthetic m17 kDa target was synthesized with an NH_2-terminal biotin. The protein was made both with and without the sequence in bold that contains the putative transmembrane sequence. Synthetic proteins were bound to streptavidin-agarose, and "catch experiments" were conducted, as described in text, using mitochondrial fractions from rat brain, skeletal muscle, or liver. Proteins eluted by reduced pH were subjected to electrophoresis on a 10% polyacrylamide gel under reducing conditions and silver stained. Lanes (−) were from beads incubated without the biotinylated protein. Lanes (+) were from beads preincubated with synthetic peptide shown. Similar results were obtained when biotinylated protein without the transmembrane sequence was used. Protein bands were cut from the gel and identified as described in text.

DISCUSSION

Although thiazolidinediones are useful in treating type 2 diabetes and have been increasingly used in the treatment of the disease (1, 47), the exact mechanism of action of these compounds remains to be defined (28). As outlined in the introduction, it is generally assumed that the overall pharmacology is secondary to direct activation of the nuclear receptor PPARγ (31, 34, 40, 43, 48). It has been hypothesized that some effects may be secondary to direct activation of PPARγ in adipose tissue; however, the pharmacology is not necessarily dependent on adipose tissue (4). As we discussed, there are some additional complexities unexplained by the hypothesis that these compounds are simple activators of PPARγ. Fur-

Table 1. *Proteins eluted from solid phase mitoNEET after binding of soluble liver mitochondrial fraction*

Complex 3 components
Ubiquinol-cytochrome *c* reductase core protein
Ubiquinol-cytochrome *c* reductase complex core protein-2
NADH dehydrogenase
NADH-ubiquinone oxidoreductase 24-kDa subunit
ATP synthase
ATP synthase α-chain
ATP synthase β-subunit
Chain G, ATPase
ATP synthase O subunit
ADP-ATP translocator
Fatty acid oxidation
Trifunctional enzyme: (gi18677763)
Long-chain acyl-CoA synthase 2
Propionyl-CoA carboxylase
PDH/SDH
Dihydrolipomide acetyltransferase component of pyruvate dehydrogenase
Dihydrolipomide succinyltransferase component of 2-oxoglutarate dehydrogenase
Succinate dehydrogenase complex, subunit A

Protein bands excised from an experiment such as that shown in Fig. 8 were analyzed text. A representative analysis from a rat liver mitochondrial sample is shown.

thermore, other direct actions of the compounds have been reported, such as inhibition of an acyl-CoA synthase (25) and direct effects on skeletal muscle (3). Several groups have suggested that thiazolidinediones may have different effects in different tissues (13, 51). Resolution of the key site of thiazolidinedione interaction that leads to useful pharmacology would undoubtedly help in the production of improved therapeutic agents.

We found that high specific activity [^{3}H]pioglitazone bound to a saturable site with relatively high affinity in the mitochondria. Active, but not inactive, analogs competed for the binding. We could not show specific [^{3}H]pioglitazone binding in any cytosolic or nuclear fractions. Because isolation of the [^{3}H]pioglitazone-binding protein(s) proved difficult, we designed a photoaffinity ligand that could be iodinated to allow purification of the protein responsible for drug binding. We have previously had success with photoaffinity ligands to locate the binding of drugs with uncharacterized binding sites (9). A single protein was cross-linked in a specific fashion in the same tissue fractions that bound [^{3}H]pioglitazone. This protein (m17) was enriched in the same subcellular fractions as [^{3}H]pioglitazone binding and demonstrated the same specificity with respect to competition by unlabeled analogs (Fig. 3).

Two approaches were taken to purify the protein from two tissue sources, bovine brain and rat liver. One involved selective solubilization with Triton X-114, followed by reverse-phase HPLC and SDS-PAGE electrophoresis. The second approach involved preparative scale electrophoresis followed by elution of the cross-linked protein from the unfixed gels, CNBr cleavage, and then electrophoresis and sequencing of the labeled fragment. The partially purified fractions submitted to trypsin digestion and MS/MS analysis, as described in MATERIALS AND METHODS, identified two common tryptic peptides from both tissue sources on three different occasions (Fig. 5). The identification was confirmed by NH_2-terminal sequencing of the 6-kDa CNBr fragment produced after preparative scale electrophoresis (Fig. 5*B*).

To provide more evidence for support of the identification, antibodies were generated to allow measurement of the protein by Western blot. In support of the findings, the antibodies recognized a protein of the same size in the same tissue fractions. Importantly, both the protein and cross-linking were undetectable under standard conditions in cultured preadipocytes, and both were measurable upon differentiation of the cells into adipocytes. The common detection under these conditions supports the conclusion that we have identified the mitochondrial protein that is cross-linked by the thiazolidinedione probe. The increased expression upon differentiation of adipocytes suggests that the protein may be involved in lipid metabolism.

Because we were unable to immunoprecipitate the cross-linked protein with the anti-peptide antibodies, we asked whether the synthetic protein would bind to solubilized mitochondrial proteins. Chromatography of soluble binding studies with tritiated pioglitazone had indicated that the mitochondrial target was associated with a large number of proteins. The identification of these proteins might provide insight as to function. The synthetic target protein bound a set of mitochondrial proteins, including components of complex 3, ATP synthase, and pyruvate dehydrogenase. Of particular interest was the association of key enzymes involved in fatty acid oxidation, especially the trifunctional enzyme complex (11). On the basis of its unique sequence and mitochondrial location, we suggest the name mitoNEET for this previously unidentified target for the thiazolidinediones. MitoNEET may serve as a supporting framework upon which mitochondrial metabolism can be channeled and controlled.

It is well known that insulin resistance is associated with elevated tissue levels of long-chain acyl-CoA esters (10). Furthermore, treatment of animals with pioglitazone can lower the tissue levels of long-chain acyl-CoA esters and even protect against their buildup during lipid infusions (49). The buildup of long-chain acyl-CoA esters could result in insulin resistance by a number of mechanisms, including activation of specific protein kinase C isoforms (50) or generation of ceramide (7). Thiazolidinediones could improve insulin sensitivity secondary to lowering the long-chain CoAs after a direct interaction with this novel mitochondrial target. Such an effect could include more efficient oxidation (11), perhaps even including a feed-forward effect by keeping malonyl-CoA levels in check (38, 39).

In conclusion, a novel mitochondrial target protein has been identified for the insulin-sensitizing thiazolidinediones. This protein may play a role in regulating mitochondrial oxidation of fatty acids, and modulation of this target may be involved in the mechanism of action of these drugs. The groundwork is now laid to test this hypothesis. Because PPARγ is known to be centrally involved in the production of new fat stores (32) and induction of glycerol kinase (16), promoting the storage of fat, it is possible that elimination of direct modulation of PPARγ while maintaining the mitochondrial interaction may produce an improved therapeutic profile.

ACKNOWLEDGMENTS

We acknowledge James P. McGrath, Timothy T. Parker, Steven P. Tanis, and John A. Easter for generation of the high specific activity tritiated pioglitazone, Linda Maggiora for synthesis of PNU-101074, Cindy Jacobs and John E. Bleasdale for the 3T3-L1 preadipocytes and adipocytes, Lonnie Adams for help with two-dimensional electrophoresis, and Ilene Reardon for bioinformatics support.

Current addresses of J. R. Colca, W. G. McDonald, J. W. Leone, and W. R. Mathews: Pfizer St. Louis, 700 Chesterfield Parkway North, Chesterfield, MO 63198; of D. J. Waldon: Guilford Pharmaceuticals, 6611 Tributary St., Baltimore, MD 21224; of J. M. Lull and C. A. Bannow: Proteos, 4717 Campus Drive Innovations Center, Kalamazoo, MI 49008; and of E. T. Lund, Pfizer, 2800 Plymouth Rd., Ann Arbor, MI 48105.

REFERENCES

1. **Bailey CJ and Day C.** Thiazolidinediones today. *Br J Diab Vasc Dis* 1: 7–13, 2001.
2. **Boyle PJ, King AB, Olansky L, Marchetti A, Lau H, Magar R, and Martin J.** Effects of pioglitazone and rosiglitazone on blood lipid levels and glycemic control in patients with type 2 diabetes mellitus: a retrospective review of randomly selected medical records. *J Clin Ther* 24: 378–396, 2002.
3. **Brunmair B, Gras F, Neschen S, Roden M, Wagner L, Waldhausl W, and Fuernsinn C.** Direct thiazolidinedione action on isolated rat skeletal muscle fuel handling is independent of peroxisome proliferator-activated receptor-[gamma]-mediated changes in gene expression. *Diabetes* 50: 2309–2315, 2001.
4. **Burant CF, Sreenan S, Hirano KI, Tai TAC, Lohmiller J, Lukens J, Davidson NO, Ross S, and Graves RA.** Troglitazone action is independent of adipose tissue. *J Clin Invest* 100: 2900–2908, 1997.
5. **Chakrabarti R, Vikramadithyan RK, Dileepkumar T, Kumar KB, Kumar MP, Misra P, Rao PB, Lohray BB, and Rajagopalan R.** Studies on the euglycemic and hypolipidemic potentials of the novel indole analogue of thiazolidinedione, DRF 2189. *Arzneimittel-Forschung* 49: 905–911, 1999.
6. **Chang AY, Wyse BM, Gilchrist BJ, Peterson T, and Diani AR.** Ciglitizone, a new hypoglycemic agent. I. Studies in ob/ob and db/db

mice, diabetic Chinese hamsters, and normal and streptozotocin-diabetic rats. *Diabetes* 32: 830–838, 1983.

7. **Chavez JA, Knotts TA, Wang LP, Li G, Dobrowsky RT, Florant GL, and Summers SA.** A role for ceramide, but not diacylglycerol, in the antagonism of insulin signal transduction by saturated fatty acids. *J Biol Chem* 278: 10297–10303, 2003.
8. **Colca JR, Dailey CF, Palazuk BJ, Hillman RM, Dinh DM, Melchior GW, and Spilman CH.** Pioglitazone hydrochloride inhibits cholesterol aborption and lowers plasma cholesterol concentrations in cholesterol-fed rats secondary to improved insulin sensitivity. *Diabetes* 40: 1669–1674, 1991.
9. **Colca JR, McDonald WG, Waldon DJ, Thomasco LM, Gadwood RG, Lund ET, Cavey GS, Mathews WR, Adams LD, Cecil ET, Pearson JD, Bock JH, Mott JE, Shinabarger DL, Xiong L, and Mankin AS.** Cross-linking in the living cell locates the site of action of oxazolidinone antibiotics. *J Biol Chem* 278: 21972–21979, 2003.
10. **Cooney GJ, Thompson AL, Furler SM, Ye J, and Kraegen EW.** Muscle long-chain acyl CoA esters and insulin resistance. *Ann NY Acad Sci* 967: 196–207, 2002.
11. **Eaton S, Bursby T, Middleton B, Pourfarzam M, Mills K, Johnson AW, and Bartlett K.** The mitochondrial trifunctional protein: center of a β-oxidation metabolon? *Biochem Soc Trans* 28: 177–182, 2000.
12. **Fleischer S and Fleischer B.** Removal and binding of polar lipids in mitochondria and other membrane systems. In: *Methods in Enzymology*, edited by Estabrook RW and Pullman ME. New York: Academic, vol. X, 1967, p. 406–433.
13. **Fuernsinn C and Waldhaeusl W.** Thiazolidinediones: metabolic actions in vitro. *Diabetologia* 45: 1211–1223, 2002.
14. **Fujita T, Sugiyama Y, Taketomi S, Sohda T, Kawamatsu Y, Iwatsuka H, and Suzuoki Z.** Reduction of insulin resistance in obese and/or diabetic animals by 5-[4-(1-methylcyclohexylmethoxy)benzyl]-thiazolidine-2,4-dione (ADD-3878, U-63287, Ciglitazone), a new antidiabetic agent. *Diabetes* 32: 804–810, 1983.
15. **Gegick CG and Altheimer MD.** Comparison of effects of thiazolidinediones on cardiovascular risk factors: observations from a clinical practice. *Endocr Pract* 7: 162–169, 2001. (Erratum: *Endocr Pract* 7: July-August 2001, p. 330)
16. **Guan HP, Li Y, Jensen MV, Newgard CB, Steppan CM, and Lazar MA.** A futile metabolic cycle activated in adipocytes by antidiabetic agents. *Nat Med* 8: 1122–1128, 2002.
17. **Hara K, Okada T, Tobe K, Yasuda K, Mori Y, Kadowaki H, Hagura R, Akanuma Y, Kimura S, Ito C, and Kadowaki T.** The Pro12Ala polymorphism in PPAR 2 may confer resistance to type 2 diabetes. *Biochem Bipohys Res Commun* 271: 212–216, 2000.
18. **Harris PK and Kletzien RF.** Localization of a pioglitazone response element in the adipocyte fatty acid-binding protein gene. *Mol Pharmacol* 45: 439–445, 1994.
19. **Hasstedt SJ, Ren QF, Teng K, and Elbein SC.** Effect of the peroxisome proliferator-activated receptor-2 Pro12Ala variant on obesity, glucose homeostasis, and blood pressure in members of familial type 2 diabetic kindreds. *J Clin Endocrinol Metab* 86: 536–541, 2001.
20. **Hiragun A, Sato M, and Mitsui H.** Preadipocyte differentiation in vitro: identification of a highly active adipogenic agent. *J Cell Physiol* 13: 124–130, 1988.
21. **Hofmann CA and Colca JR.** New oral thiazolidinedione antidiabetic agents act as insulin sensitizers. *Diabetes Care* 15: 1075–1078, 1992.
22. **Ikeda H, Taketomi S, Sugiyama Y, Shimura Y, Sohda T, and Meguro K.** Effects of pioglitazone on glucose and lipid metabolism in normal and insulin resistant animals. *Arzneimittel-Forschung* 40: 156–162, 1990.
23. **Ikonen E, Fiedler K, Parton RG, and Simons K.** Prohibitin, an antiproliferative protein, is localized to mitochondria. *FEBS Lett* 358: 273–277, 1995.
24. **Khan MA, St. Peter JV, and Xue JL.** A prospective, randomized comparison of the metabolic effects of pioglitazone or rosiglitazone in patients with type 2 diabetes who were previously treated with troglitazone. *Diabetes Care* 25: 708–711, 2002.
25. **Kim JH, Lewin TM, and Coleman RA.** Expression and characterization of recombinant rat acyl-CoA synthetases 1, 4, and 5. Selective inhibition by triacsin C and thiazolidinediones. *J Biol Chem* 276: 24667–24673, 2001.
26. **Kletzien RF, Clarke SD, and Ulrich RG.** Enhancement of adipocyte differentiation by an insulin-sensitizing agent. *Mol Pharmacol* 41: 393–398, 1992.
27. **Kletzien RF, Foellmi LA, Harris PK, Wyse BM, and Clarke SD.** Adipocyte fatty acid-binding protein: regulation of gene expression in vivo and in vitro by an insulin-sensitizing agent. *Mol Pharmacol* 42: 558–562, 1992.
28. **Komers R and Vrana A.** Thiazolidinediones—tools for the research of metabolic syndrome X. *Physiol Res* 47: 215–225, 1998.
29. **LaCivita KA and Villarreal G.** Differences in lipid profiles of patients given rosiglitazone followed by pioglitazone. *Curr Med Res Opin* 18: 363–370, 2002.
30. **Lehmann JM, Lenhard JM, Oliver BB, Ringold GM, and Kliewer SA.** Peroxisome proliferator-activated receptors α and γ are activated by indomethacin and other non-steroidal anti-inflammatory drugs. *J Biol Chem* 272: 3406–3410, 1997.
31. **Lehmann JM, Moore LB, Smith-Oliver TA, Wilkison WO, Willson TM, and Kliewer SA.** An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor [gamma] (PPAR [gamma]). *J Biol Chem* 270: 12953–12956, 1995.
32. **Lowell BB.** An essential regulator of adipogenesis and modulator of fat cell function. *Cell* 99: 239–242, 1999.
33. **Miles PD, Barak Y, He W, Evans RM, and Olefsky JM.** Improved insulin-sensitivity in mice heterozygous for PPAR-gamma deficiency. *J Clin Invest* 105: 287–292, 2000.
34. **Olefsky JM and Saltiel AR.** PPAR gamma and the treatment of insulin resistance. *Trends Endocrinol Metab* 11: 362–368, 2000.
35. **Owens DR.** Thiazolidinediones: a pharmacological overview. *Clin Drug Invest* 22: 485–505, 2002.
36. **Rami HK and Smith SA.** Synthetic ligands for PPARγ—review of patent literature 1994–1999. *Exp Opin Ther Patents* 10: 623–634, 2000.
37. **Reginato MJ, Bailey ST, Krakow SL, Minami CO, Ishii S, Tanaka H, and Lazar MA.** A potent antidiabetic thiazolidinedione with unique peroxisome proliferator-activated receptor-activating properties. *J Biol Chem* 273: 32679–32684, 1998.
38. **Ruderman NB, Saha AK, Vavvas D, and Witters LA.** Malonyl-CoA, fuel sensing, and insulin resistance. *Am J Physiol Endocrinol Metab* 276: E1–E18, 1999.
39. **Saha AK, Kurowski TG, Colca JR, and Ruderman NB.** Lipid abnormalities in tissues of the KKAy mouse: effects of pioglitazone on malonyl-CoA and diacylglycerol. *Am J Physiol Endocrinol Metab* 267: E95–E101, 1994.
40. **Saltiel AR and Olefsky JM.** Thiazolidinediones in the treatment of insulin resistance and type II diabetes. *Diabetes* 45: 1661–1669, 1996.
41. **Sizer KM, Smith CL, Jacob CS, Swanson ML, and Bleasdale JE.** Pioglitazone promotes insulin-induced activation of phosphoinositide 3-kinase in 3T3-L1 adipocytes by inhibiting a negative control mechanism. *Mol Cell Endocrinol* 102: 119–129, 1994.
42. **Sorbera LA, Castaner J, Del Fresno M, and Silvestre J.** Netoglitazone. In: *Drugs of the Future* 27: 132–139, 2002.
43. **Spiegelman BM.** PPAR-[gamma]: adipogenic regulator and thiazolidinedione receptor. *Diabetes* 47: 507–514, 1998.
44. **Stumvoll M and Haring H.** The peroxisome proliferator-activated receptor γ2-Pro12Ala polymorphism. *Diabetes* 51: 2341–2347, 2002.
45. **Tanis SP, Parker TT, Colca JR, Fisher RM, and Kletzien RF.** Synthesis and biological activity of metabolites of the antidiabetic, antihyperglycemic agent pioglitazone. *J Med Chem* 39: 5053–5063, 1996.
46. **Uppenberg J, Svensson C, Jaki M, Bertilsson G, Jendeberg L, and Berkenstam A.** Crystal structure of the ligand binding domain of the human nuclear receptor PPARgamma. *J Biol Chem* 273: 31108–31112, 1998.
47. **Wagstaff AJ and Goa KL.** Rosiglitazone: a review of its use in the management of type 2 diabetes mellitus. *Drugs* 62: 1805–1837, 2002.
48. **Willson TM, Cobb JE, Cowan DJ, Wiethe RW, Correa ID, Prakash SR, Beck KD, Moore LB, Kliewer SA, and Lehmann JM.** The structure-activity relationship between peroxisome proliferator-activated receptor [gamma] agonism and the antihyperglycemic activity of thiazolidinediones. *J Med Chem* 39: 665–668, 1996.
49. **Ye JM, Frangioudakis G, Iglesias MA, Furler SM, Ellis B, Dzamko N, Cooney GJ, and Kraegen EW.** Prior thiazolidinedione treatment preserves insulin sensitivity in normal rats during acute fatty acid elevation: role of the liver. *Endocrinology* 143: 4527–4535, 2002.
50. **Yu C, Chen Y, Cline GW, Zhang D, Zong H, Wang Y, Bergeron R, and Kim JK, Cushman SW, Cooney GJ, Atcheson B, White MF, Kraegen EW, and Shulman GI.** Mechanism by which fatty acids inhibit insulin activation of insulin receptor substrate-1 (IRS-1)-associated phosphatidylinositol 3-kinase activity in muscle. *J Biol Chem* 277: 50230–50236, 2002.
51. **Zierath JR, Ryder JW, Doebber T, Woods J, Wu M, Ventre J, Li Z, McCrary C, Berger J, Zhang B, and Moller DE.** Role of skeletal muscle in thiazolidinedione insulin sensitizer (PPAR agonist) action. *Endocrinology* 139: 5034–5041, 1998.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.